



# **STIC Search Report**

## **Biotech-Chem Library**

STIC Database Tracking Number: 105234

**TO: Phillip Gambel**  
**Location: CM1/8B03&9E12**  
**Art Unit: 1644**  
**Wednesday, October 08, 2003**

**Case Serial Number: 09/730466**

**From: Toby Port**  
**Location: Biotech-Chem Library**  
**CM1-6A04**  
**Phone: 308-3534**

**toby.port@uspto.gov**

### **Search Notes**

Dear Examiner Gambel,

Here are the results of your search.  
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

un on: October 7, 2003, 11:09:47 ; Search time 83 Seconds

(without alignments)  
175.938 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_1\_92

Perfect score: 475

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....AFSFKNRVYLDVTYSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475	100.0	112	23	Human transmembran
2	475	100.0	120	13	LFA-3 CD2 binding
3	475	100.0	128	17	Human cell adhesio
4	475	100.0	133	17	Human cell adhesio
5	475	100.0	133	17	Human cell adhesio
6	475	100.0	237	13	Human LFA-3 antige
7	475	100.0	237	17	Human LFA-3 antige
8	475	100.0	237	19	Human LFA-3 antige
9	475	100.0	237	20	Human LFA-3 antige

10	475	100.0	237	21	AA196127	Human cell surface
11	475	100.0	237	22	AAU02436	Human lymphocyte f
12	475	100.0	240	11	AAR05572	PI-linked LFA-3 em
13	475	100.0	240	11	AAR07604	Lymphocyte functio
14	475	100.0	240	13	AAR27162	PI-linked LFA-3 us
15	475	100.0	240	21	AA183134	PI-linked human tr
16	475	100.0	240	23	AA181158	Human PI-linked tr
17	475	100.0	240	23	AAU76226	PI-linked lymphocy
18	475	100.0	250	9	AA181507	Human LFA-3 expres
19	475	100.0	250	13	AA182161	Transmembrane LFA-
20	475	100.0	250	14	AA184371	Sequence encoded b
21	475	100.0	250	15	AA186471	Human LFA-3. Homo
22	475	100.0	250	17	AA184370	Human cell adhesio
23	475	100.0	250	21	AA183133	Human transmembran
24	475	100.0	250	22	AA181157	Human transmembran
25	475	100.0	250	23	AAU76225	Lymphocyte functio
26	475	100.0	256	24	AAU76225	Murine CD2-binding
27	475	100.0	347	13	AA1827163	CD2 binding LFA-3-
28	475	100.0	347	21	AA183136	Human transmembran
29	475	100.0	347	22	AA183136	Human LFA3TIP fusi
30	475	100.0	347	23	AA181160	Human lymphocyte f
31	475	100.0	347	23	ABG70766	Human LFA3TIP poly
32	475	100.0	347	23	AB181994	LFA3/IgG fusion pr
33	475	100.0	347	23	AAU76228	Murine LFA-3. Mus
34	475	100.0	351	24	AB181995	Human LFA-3/IgG fu
35	472	99.4	240	14	AA184222	Amino acid sequenc
36	472	99.4	240	14	AA184222	Sequence encoded b
37	472	99.4	250	14	AA184221	Amino acid sequenc
38	469	98.7	237	18	AA181687	Human CD58 GPI. H
39	464	97.7	237	12	AA181482	Human LFA-3(CD58)
40	463	97.5	347	14	AA184224	Amino acid sequenc
41	459	96.6	250	13	AA1828365	Human LFA-3 protei
42	443	93.3	280	22	AB181595	Novel human diagn
43	328	69.1	71	13	AA1827158	LFA-3 CD2 binding
44	328	69.1	134	13	AA1828364	Human LFA-3 D2 reg
45	328	69.1	135	13	AA1828369	Human LFA-3-delta-

ALIGNMENTS

RESULT 1  
ID ABB81992 standard; Protein; 112 AA.  
XX ABB81992;  
XX

25-NOV-2002 (first entry)

Human transmembrane LFA-3 partial polypeptide.

Human; immunosuppressive; antitumor; dermatological; antiinflammatory;  
antiproliferative; antineoplastic; antitumor; antitumor; antitumor;  
antiproliferative; antineoplastic; antitumor; antitumor; antitumor;  
osteoporotic; antidiabetic; gene therapy; CD2; LFA-3; interferon; IFN;  
CD25; CD8; CD45.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..28

FT Protein /note= "signal peptide"

FT /note= "partial mature protein fragment; the sequence

after position 84 is not indicated in the  
specification; the pages containing the complete  
sequence are not provided"

WO20020480-A1.

08-AUG-2002.

25-JAN-2002; 2002WO-US02314.

XX

R 01-FEB-2001; 2001US-265964P.

X (BIOJ ) BIOGEN INC.

X Vaishnav AK, Cooper KD, Shrager D, McCormick TS;

X WPI; 2002-657488/70.

R N-PSDB; ABQ79655.

X Treating or preventing, by inhibiting CD2/LFA-3 interaction, an  
T epidermal or dermal disorder, inflammatory and/or autoimmune disorder  
T with aberrant T cell activity or proliferation -

X Disclosure; Page 58; 68pp; English.

X The invention relates to treating or preventing an epidermal or dermal  
C disorder, an inflammatory disorder and/or an autoimmune disorder in a  
C subject with aberrant T cell activity or proliferation. The method  
C involves administering an inhibitor of the CD2/LFA-3 interaction, in  
C combination with an auxiliary agent, thereby treating or preventing the  
C epidermal or dermal disorder, the inflammatory disorder or the autoimmune  
C disorder. The methods and compositions of the invention are useful for  
C preventing or treating skin disorders characterized by increased T cell  
C activation and abnormal antigen presentation in the dermis and epidermis,  
C such as psoriasis, UV damage, atopic dermatitis, cutaneous T cell  
C lymphoma, contact and atopic dermatitis, lichen planus, alopecia,  
C vitiligo, urticaria. They may also be used in chronic inflammatory and  
C autoimmune disorders such as diabetes mellitus, arthritis, rheumatoid  
C arthritis, juvenile rheumatoid arthritis, osteoarthritis, psoriatic  
C arthritis, multiple sclerosis, encephalomyelitis, myasthenia gravis,  
C systemic lupus erythematosus, autoimmune thyroiditis. The present  
C sequence represents the human transmembrane LFA-3 partial amino acid  
C sequence.

SQ Sequence 112 AA;

Query Match 100.0%; Score 475; DB 23; Length 112;

Best Local Similarity 100.0%; Pred. No. 4.5e-51;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFIQSCFQIYGVYGVNTFHVPSNPLKEVLKKQK 60

DB 1 MVAGSDAGRALGVLSVCLLHCFGFIQSCFQIYGVYGVNTFHVPSNPLKEVLKKQK 60

QY 61 DKVAELENSEFRASFKNRYLDTVSGSLTI 92

DB 61 DKVAELENSEFRASFKNRYLDTVSGSLTI 92

RESULT 2

AAAR27157

ID AAR27157 standard; protein; 120 AA.

XX

XX

AC AAR27157;

DT 25-MAR-2003 (updated)

DT 20-MAR-1998 (first entry)

XX

DE LFA-3 CD2 binding domain polypeptide #2.

XX

KW Lymphocyte associated antigen-3; T-lymphocyte accessory molecule.

XX

OS Synthetic.

XX

Key Location/Qualifiers

FT region 2..78

FT /note= "1-77 of these residues may be deleted"

FT region 89..120

FT /note= "1-32 of these residues may be deleted"

XX EF503648-A1.

XX

PD 16-SEP-1992.

XX 12-MAR-1992; 92EP-0104320.

XX 12-MAR-1991; 91US-0667971.

PR 07-OCT-1991; 91US-0770967.

XX (BIOJ ) BIOGEN INC.

PI Miller GT, Rosa MD, Wallner BP;

XX WPI; 1992-309760/38.

DR CD2-binding domain of lymphocyte function associated antigen-3

PT and DNA - for diagnosing and treating inflammation and

PT autoimmune diseases, e.g. systemic lupus erythematosus and

PT rheumatoid arthritis

XX Claim 1; Page 52-53; 85pp; English.

XX This polypeptide can bind to CD2. The N-terminal Met is opt.

CC absent. The N-terminal and C-terminal portions can be deleted.

CC The polypeptide and its functional deletion mutants may be used to

CC treat acute and chronic inflammation, autoimmune disease and for

CC immunomodulation. The polypeptides can also form the N-terminal

CC part of a fusion protein. The polypeptides and fusion proteins may

CC also be used to inhibit T-cell activation and the proliferation of

CC peripheral blood lymphocytes. Multimeric proteins can be formed from

CC the polypeptides and/or fusion proteins. The multimers have enhanced

CC affinity for CD2. See also AAR27157-9.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 120 AA;

Query Match 100.0%; Score 475; DB 13; Length 120;

Best Local Similarity 100.0%; Pred. No. 4.9e-51;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFIQSCFQIYGVYGVNTFHVPSNPLKEVLKKQK 60

DB 1 MVAGSDAGRALGVLSVCLLHCFGFIQSCFQIYGVYGVNTFHVPSNPLKEVLKKQK 60

QY 61 DKVAELENSEFRASFKNRYLDTVSGSLTI 92

DB 61 DKVAELENSEFRASFKNRYLDTVSGSLTI 92

RESULT 3

AAW04361

ID AAW04361 standard; Protein; 128 AA.

XX

XX

AC AAW04361;

XX

DT 27-JUN-1997 (first entry)

XX

DE Human cell adhesion protein LFA-3.

XX

KW Human; T cell; MOLT-4; cell adhesion; LFA-3; immunosuppression;

XX cell line; immunosuppressant cell.

XX

OS Homo sapiens.

XX

PN W09633217-A1.

XX

PD 24-OCT-1996.

XX

XX 15-APR-1996; 96WO-JP01039.

XX

PR 27-DEC-1995; 95JP-0341959.

PR 19-APR-1995; 95JP-0094060.

XX

XX 04-JUL-1995; 95JP-0169110.

XX

PA (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.

XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: October 7, 2003, 11:13:47 ; Search time 40 seconds  
(without alignments)  
221.188 Million cell updates/sec

Title: us-09-730-465-2\_copy\_1\_92

Perfect score: 475

Sequence: 1 MVAGSDAGRALGVLSVCLL.....AFSSFKNRVLDFTVSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76.\*

2: PIR1.\*

3: PIR2.\*

4: PIR3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	475	100.0	250	2 A28564	lymphocyte functio
2	74.5	15.7	243	2 A53244	leukocyte antigen
3	70	14.7	546	2 T25568	hypothetical prote
4	69.5	14.6	389	2 T16326	hypothetical prote
5	67.5	14.2	240	2 JL0143	antigen BCM1 precu
6	66	13.9	661	2 S49901	coat protein gpl -
7	66	13.9	707	2 JC7763	neuronal leucine-r
8	65	13.7	416	2 T25036	hypothetical prote
9	65	13.7	1822	2 S44849	K12H4.8 protein -
10	64	13.5	737	2 T13504	NADH2 dehydrogenas
11	63.5	13.4	562	2 C82355	hypothetical prote
12	63	13.3	294	2 AD1649	weakly phage relat
13	63	13.3	299	2 E90162	conserved hypothe
14	63	13.3	608	2 S58091	probable membrane
15	62.5	13.2	274	2 S76154	hypothetical prote
16	62.5	13.2	535	2 T23874	hypothetical prote
17	62.5	13.2	734	2 T13685	NADH2 dehydrogenas
18	62	13.1	237	2 C70157	hypothetical prote
19	62	13.1	403	2 C33958	hypothetical prote
20	62	13.1	414	2 H84936	tRNA adenylytrans
21	62	13.1	426	2 A81933	probable malate de
22	62	13.1	426	2 C81173	malate oxidoreduct
23	62	13.1	510	2 E70321	GMP synthase (glut
24	61.5	12.9	522	2 E82024	peptide methionine
25	61.5	12.9	522	2 G81243	peptide methionine
26	61.5	12.9	741	2 T12711	NADH2 dehydrogenas
27	61.5	12.9	752	2 E97191	uncharacterized pr
28	61.5	12.9	836	2 E97851	hypothetical prote
29	61	12.8	344	2 T00993	probable beta-1,3-

30 61 12.8 347 2 S41638  
31 61 12.8 405 2 S41917  
32 61 12.8 424 1 XNBYO  
33 61 12.8 508 1 A30007  
34 61 12.8 539 2 T22041  
35 61 12.8 1107 1 S52517  
36 60.5 12.7 275 2 E83655  
37 60.5 12.7 351 2 T40762  
38 60.5 12.7 380 2 T17423  
39 60.5 12.7 475 2 T33979  
40 60.5 12.7 490 2 T38088  
41 60.5 12.7 1210 2 C59431  
42 60.5 12.7 2924 2 T18378  
43 60 12.6 496 2 JH0165  
44 60 12.6 496 2 S46459  
45 60 12.6 526 2 JC7248

## ALIGNMENTS

## RESULT 1

A28564  
Lymphocyte function-associated antigen 3, transmembrane splice form precursor - huma  
N:Alternate names: CD58 antigen, surface glycoprotein LFA-3  
C:Species: Homo sapiens (man)  
C>Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 02-Aug-2002  
C:Accession: A28564; S01269  
R:Wallner, B.P.; Frey, A.Z.; Tizard, R.; Mattaliano, R.J.; Hession, C.; Sanders, M.F  
J. Exp. Med. 166, 923-932, 1987  
A:Title: Primary structure of lymphocyte function-associated antigen 3 (LFA-3). The  
A:Reference number: A28564; MUID:88009714; PMID:3309127  
A:Accession: A28564  
A:Molecule type: mRNA  
A:Residues: 1-250 <WAL>  
A:Cross-references: GB:Y00636; NID:g34346; PIDN:CAA68668.1; PID:g34347  
A:Experimental source: erythrocytes  
A:Note: the authors translated the codon TAT for residue 34 as Val  
R:Seed, B.  
Nature 329, 840-842, 1987  
A:Title: An LFA-3 CDNA encodes a phospholipid-linked membrane protein homologous to  
A:Reference number: S01269; MUID:88039074; PMID:3313052  
A:Accession: S01269  
A:Molecule type: mRNA  
A:Residues: 1-235, 'VL' <SEE>  
A:Cross-references: EMBL:X06296; NID:g34349; PIDN:CRA29622.1; PID:g34350  
C:Comment: For an alternative splice form, see PIR:S01269.  
C:Comment: For an alternative splice form, see PIR:A28564.  
C:Genetics:  
A:Gene: GDB:CD58; LFA3  
A:Cross-references: GDB:120580; OMIM:153420  
A:Map position: lpi3-lpi3  
C:Superfamily: human B-cell antigen CD19; immunoglobulin homology  
C:Keywords: alternative splicing; glycoprotein; phosphatidylinositol linkage; surfac  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-250/Product: lymphocyte function-associated antigen 3 transmembrane splice form  
F:135-189/Domain: immunoglobulin homology <IMM>  
F:40,94,109,135,169,195/Binding site: carbohydrate (Asn) (covalent) #status predicts

Query Match 100.0%; Score 475; DB 2; Length 250;

Best Local Similarity 100.0%; Pred. No. 2.3e-45; Mismatches 0; Indels 0; Gaps 0;

Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFCSCFSQQIYGVYGVNVTFFHVPNNVPLKVLKQK 60

Db 1 MVAGSDAGRALGVLSVCLLHCFGFCSCFSQQIYGVYGVNVTFFHVPNNVPLKVLKQK 60

QY 61 DKVAELENSEFPASFNNRVYLDFTVSGSLTI 92

Db 61 DKVAELENSEFPASFNNRVYLDFTVSGSLTI 92

## RESULT 2

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C/Accession: T26568

R;White, S.

submitted to the EMBL Data Library, September 1999

A/Reference number: Z20234

A/Accession: T26568

A/Status: preliminary; translated from GE/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-546 <WILL>

A/Cross-references: EMBL:AL110478; NID:g1542139; PIDN:CAB54346.1; CESP:Y26D4A.10

A/Experimental source: clone Y26D4A

C/Genetics:

A/Gene: CESP:Y26D4A.10

A/Introns: 415/3; 506/3

Query Match 14.7%; Score 70; DB 2; Length 546;  
Best Local Similarity 28.9%; Pred. No. 8.2; 27; Indels 16; Gaps 3;  
Matches 22; Conservative 11; Mismatches

QY 19 LLHCFGFI SCFOIYG VVY-----GNVTFFVSPNVLKVKLMKKQ--KDKV-- 63  
| | | | :  
Db 163 LTHIFGVCEYKIIYGRKNLSLYRSPENKNSITSSNPONFTAKRVINKLWIIPKIIP 222

QY 64 -AELENSEFRAFSSEFK 78  
| : : : | | | | |

Db 223 CLEMQNKLLNIFESNK 238

RESULT 4

Tl6326

hypothetical protein F41C6.7 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C/Accession: Tl6326

R;Geisel, C.

submitted to the EMBL Data Library, October 1995

A/Description: The sequence of C. elegans cosmid F41C6.

A/Reference number: Z18495

A/Accession: Tl6326

A/Status: preliminary; translated from GE/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-389 <GEI>

A/Cross-references: EMBL:U39745; NID:g1049470; PID:g1049476; PIDN:AAA80448.1; CESP:F41C6.7

A/Gene: CESP:F41C6.7

A/Introns: 50/3; 118/3; 136/3; 154/3; 189/3; 209/3; 235/2; 264/2; 300/3; 327/3; 327/3; 362/3

Query Match 14.6%; Score 69.5; DB 2; Length 389;  
Best Local Similarity 28.3%; Pred. No. 6.4;  
Matches 17; Conservative 14; Mismatches 28; Indels 1; Gaps 1;

QY 14 LSIVCLLHCFGFI SCFOIYG VVYGNVTFFVSPNVLKVKLMKKQKDYAELENSEFRA 73  
| | | :  
Db 289 LVAVCIASHYWKLIISNQSMFTVIGYTIHV-GNKAIVEIYEVENQQNRREIQSFSA 347

RESULT 5

JL0143

antigen BCM1 precursor - mouse

N/Alternate names: CD48 antigen homolog sgp-60; OX45 antigen, Blast-1 antigen

C/Species: Mus musculus (house mouse)

C/Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 31-Jan-2000

C/Accession: JL0143; S21319; A47469; B47469

R;Wong, Y.W.; Williams, A.F.; Kingmsore, S.F.; Seidin, M.F.  
J. Exp. Med. 171, 2115-2130, 1990

A>Title: Structure, expression, and genetic linkage of the mouse BCM1 (OX45 or Blast-1) region on mouse chromosome 3.

A/Reference number: JL0143; MID:90278362; PMID:1693656

A/Accession: JL0143

A/Molecule type: mRNA

A/Residues: 1-240 <WON>

A/Cross-references: EMBL:X17501; NID:g50134; PIDN:CAA35542.1; PID:g50135

R;Wong, Y.W.; Williams, A.F.; Kingmsore, S.F.; Seidin, M.F.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:12:57 ; Search time 22 Seconds  
(without alignments)  
196.657 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_1\_92

Perfect score: 475

Sequence: 1 MVAGSDAGRALGVLSVVCIL.....AFSFKNRVLDVSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	475	100.0	250	1 LFA3_HUMAN	P19256 homo sapien
2	74.5	15.7	243	1 CD48_HUMAN	P09326 homo sapien
3	67.5	14.2	240	1 CD48_MOUSE	P18181 mus musculus
4	63.5	13.4	611	1 GLMS_WIGER	Q8d3j0 w glucosami
5	63	13.3	379	1 ASPN_HUMAN	Q9bxx1 homo sapien
6	62	13.1	414	1 CCA_BUCAI	P57169 buchnera ap
7	62	13.1	461	1 PUCQ_RHOCA	P23462 rhodobacter
8	62	13.1	510	1 GUAA_AOUEA	O66601 aquifex aeo
9	62	13.1	552	1 YHJL_SCHPO	Q9c0v8 schizosacch
10	61.5	12.9	522	1 MSAB_NEIGO	P14930 n peptide m
11	61.5	12.9	522	1 MSAB_NEIMA	Q9jwm8 n peptide m
12	61.5	12.9	522	1 MSAB_NEIMB	Q9kin8 n peptide m
13	61.5	12.9	741	1 NUSC_CICIN	Q32007 cichorium i
14	61	12.8	347	1 CD2_HORSE	P37998 equus cabal
15	61	12.8	373	1 ASPN_MOUSE	Q9p9g4 mus musculu
16	61	12.8	408	1 METK_DROME	P40320 drosophila
17	61	12.8	424	1 ORAT_FEAST	P07991 saccharomyc
18	61	12.8	508	1 GSPB_CHICK	P12244 gallus gall
19	61	12.8	1107	1 MYLE_RAT	O63356 rattus norv
20	60.5	12.7	193	1 Y304_SUITO	Q975x4 sulfolobus
21	60.5	12.7	262	1 Y931_STRPY	Q9a049 streptococ
22	60.5	12.7	490	1 YAM7_SCHPO	O10662 schizosacch
23	60.5	12.7	1136	1 CRD2_HUMAN	Q96948 homo sapien
24	60	12.6	496	1 GRB_MOUSE	P48168 mus musculu
25	60	12.6	496	1 GRB_RAT	P20781 rattus norv
26	60	12.6	526	1 S17A_MOUSE	Q9qz39 mus musculu
27	59.5	12.5	609	1 RFAL_XENLA	Q01588 xenopus lae
28	59.5	12.5	741	1 NUSC_ASTCO	P51097 aster cordi
29	59.5	12.5	974	1 ATXB_LEIDO	P12522 leishmania
30	59.5	12.5	1656	1 ATC8_YEAST	Q12674 saccharomyc
31	59	12.4	268	1 ILLA_BOVIN	P08831 bos taurus
32	59	12.4	288	1 CD80_HUMAN	P33681 homo sapien
33	59	12.4	314	1 ISPH_HABIN	P44976 haemophilus

34	59	12.4	438	1 ARLY_CLOTE	P59616 clostridium
35	59	12.4	511	1 GUAA_CAMJE	Q9pnd9 campylobact
36	59	12.4	741	1 NUSC_FLARA	Q32238 flaveria ra
37	59	12.4	1109	1 MYLE_HUMAN	Q12965 homo sapien
38	58.5	12.3	165	1 GCSH_ARATH	P25855 arabidopsis
39	58.5	12.3	210	1 SODF_SULSO	P80857 sulfolobus
40	58.5	12.3	348	1 Y258_HELPV	P56136 helicobacte
41	58.5	12.3	351	1 MURG_HAEIN	P45085 haemophilus
42	58.5	12.3	447	1 ANCA_CLOFM	Q66848 clostridium
43	58.5	12.3	508	1 GUAA_HELPV	Q9zkg4 helicobacte
44	58.5	12.3	1069	1 ENTK_MOUSE	P97435 mus musculu
45	58	12.2	631	1 EMF2_ARATH	Q81674 arabidopsis

## ALIGNMENTS

RESULT 1  
LFA3\_HUMAN  
ID LFA3\_HUMAN STANDARD; PRT; 250 AA.  
AC P19256; Q96KI9;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Lymphocyte function-associated antigen 3 precursor (Ag3) (Antigen CD58) (Surface glycoprotein LFA-3).  
GN CD58 OR LFA3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
RC TISSUE=Erythrocyte;  
RX MEDLINE=88009714; PubMed=3309127;  
RA Wallner B.P., Frey A.Z., Tizard R., Mattaliano R.J., Hession C., Sanders M.E., Dustin M.L., Springer T.A.;  
RT "Primary structure of lymphocyte function-associated antigen 3 (LFA-3). The ligand of the T lymphocyte CD2 glycoprotein.";  
RL J. Exp. Med. 186:923-932(1987).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RX MEDLINE=88039074; PubMed=3313052;  
RA Seed B.;  
RT "An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2.";  
RL Nature 329:840-842(1987).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RX MEDLINE=98169008; PubMed=9510189;  
RA Wallich R., Brenner C., Brand Y., Roux M., Reister M., Meuer S.;  
RT "Gene structure, promoter characterization, and basis for alternative mRNA splicing of the human CD58 gene.";  
RL J. Immunol. 160:2862-2871(1998).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 29-122.  
RX MEDLINE=99218273; PubMed=10200255;  
RA Ikemizu S., Sparks L.M., van der Merwe P.A., Harlos K., Stuart D.I., Jones E.Y., Davis S.J.;  
RT "Crystal structure of the CD2-binding domain of CD58 (Lymphocyte function-associated antigen 3) at 1.8-A resolution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:4289-4294(1999).  
CC -1- FUNCTION: LIGAND OF THE T LYMPHOCYTE CD2 GLYCOPROTEIN. THIS INTERACTION IS IMPORTANT IN MEDIATING THYMOCYTE INTERACTIONS WITH THYMIC EPITHELIAL CELLS, ANTIGEN-INDEPENDENT AND -DEPENDENT INTERACTIONS OF T LYMPHOCYTES WITH TARGET CELLS AND ANTIGEN-PRESENTING CELLS AND THE T LYMPHOCYTE ROSETTING WITH ERYTHROCYTES. IN ADDITION, THE LFA-3/CD2 INTERACTION MAY PRIME RESPONSE BY BOTH THE CD2+ AND LFA-3+ CELLS.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and attached to the membrane by a GPI-anchor (isoform 2).  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms-2;

RESULT 2

ID	CD48_HUMAN	STANDARD;	PRT;	243 AA.
AC	P09326;			
AD	01-MAR-1989 (Rel. 10, Created)			
AT	01-NOV-1991 (Rel. 20, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	B-lymphocyte activation marker BLAST-1 precursor (BCM1 surface antigen) (Leucocyte antigen MEM-102) (TCT-1) (Antigen CD48).			
DE	CD48 OR BCM1 OR BLAST1.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OC	NCBI_TaxID=9606;			
OT	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=91153858; PubMed=1999351;			
RP	Vaughan H.A., Henning M.M., Purcell D.F.J., McKenzie I.F.C.,			
RP	Sandrin M.S.;			
RA	"The isolation of cDNA clones for CD48.";			
RA	[2]			
RP	Immunogenetics 33:113-117(1991).			
RP	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=91153857; PubMed=1999350;			
RP	Korinek V., Stefanova I., Angelisova P., Hilbert I., Horejsi V.;			
RP	"The human leucocyte antigen CD48 (MEM-102) is closely related to the			
RP	activation marker Blast-1.";			
RP	Immunogenetics 33:108-112(1991).			
RP	[3]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=88111558; PubMed=2826034;			
RP	Staunton D.E., Thorley-Lawson D.A.;			
RA	"Molecular cloning of the lymphocyte activation marker Blast-1.";			
RA	EMBO J. 6:3695-3701(1987).			
RA	[4]			
RP	SEQUENCE FROM N.A.			
RP	Bates K.;			
RA	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
RA	[5]			
RP	SEQUENCE FROM N.A.			
RP	TISSUE=B-cell;			
RP	MEDLINE=22386257; PubMed=12477932;			
RP	Strausberg R.I., Feingold E.A., Grouse L.H., Derge J.G.,			
RP	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RP	Altschul S.F., Zeeberg B., Buetkow K.H., Schaefer C.F., Bhat N.K.,			
RP	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RP	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RP	Diatchenko L., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,			
RP	Stapleton M.J., Udels T.B., Toshivuki S., Carpinci P., Prange C.,			
RP	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RP	Bosk S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RP	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RP	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RP	Fahney J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,			
RP	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RP	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RP	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RP	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smillius D.E.,			
RP	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RP	"Generation and initial analysis of more than 15,000 full-length			
RP	human and mouse cDNA sequences.";			
RP	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RP	[6]			
RP	SEQUENCE OF 1-27 FROM N.A.			
RP	MEDLINE=91141511; PubMed=1847502;			
RP	Fisher R.C., Thorley-Lawson D.A.;			
RA	"Characterization of the Epstein-Barr virus-inducible gene encoding			
RA	the human leukocyte adhesion and activation antigen BLAST-1 (CD48).";			
RA	Mol. Cell. Biol. 11:1614-1623(1991).			
RA	[7]			
RP	PARTIAL SEQUENCE.			
RP	MEDLINE=91237281; PubMed=1827826;			

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

3M protein - protein search, using sw model

Run on: October 7, 2003, 11:09:22 ; Search time 96 seconds  
(without alignments)

247.300 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_1\_92

Perfect score: 475

Sequence: 1 MVAGSDAGRALGVLSVWCLL.....AFSFKNRNYLDTVSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp.archaea:\*

2: sp.bacteria:\*

3: sp.fungi:\*

4: sp.human:\*

5: sp.invertebrate:\*

6: sp.mammal:\*

7: sp.mhc:\*

8: sp.organelle:\*

9: sp.phage:\*

10: sp.plant:\*

11: sp.podent:\*

12: sp.virus:\*

13: sp.vertibrate:\*

14: sp.unclassified:\*

15: sp.rvirus:\*

16: sp.bacteriap:\*

17: sp\_cheap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	475	100.0	240	4 Q9BRW0	Q9brw0 homo sapien
2	328	69.1	134	4 Q14748	Q14748 homo sapien
3	287	60.4	119	4 Q16393	Q16393 homo sapien
4	256.5	54.0	159	6 Q28752	Q28752 ovis sp. lf
5	256.5	54.0	227	6 Q28754	Q28754 ovis sp. lf
6	256.5	54.0	253	6 Q28753	Q28753 ovis sp. lf
7	137.5	33.2	244	6 Q8Q8B6	Q8sqb6 sus scrofa
8	122	25.7	357	13 Q90ZL5	Q90z15 anas platyr
9	76	16.0	108	2 Q8GN34	Q8gn34 uncultured
10	75.5	15.9	1225	2 Q8GAT8	Q8gat8 altercomonas
11	74.5	15.7	169	7 Q8MGRO	Q8mgr0 homo sapien
12	73	15.4	928	5 Q8IJN2	Q8ijn2 plasmodium
13	70	14.7	288	6 Q77684	Q77684 macaca neme
14	70	14.7	288	6 Q28499	Q28499 macaca mla
15	70	14.7	546	5 Q90ZP8	Q90zp8 caenorhabdi
16	69.5	14.6	389	5 Q20272	Q20272 caenorhabdi

17	68	14.3	499	3 Q9US44	Q9us44 schizosacch
18	67.5	14.2	499	5 Q9GP65	Q9gp65 trypanosoma
19	67.5	14.2	499	5 Q9GU80	Q9gu80 trypanosoma
20	67.5	14.2	946	11 Q9IY12	Q9iy12 mus musculu
21	67	14.1	200	12 Q9IW07	Q9iw07 poinsettia
22	67	14.1	1987	12 Q9IW08	Q9iw08 poinsettia
23	66.5	14.0	151	11 Q8C2T1	Q8c2t1 mus musculu
24	66.5	14.0	901	11 Q88191	Q88191 mus musculu
25	66.5	14.0	937	11 Q9IY13	Q9iy13 mus musculu
26	66	13.9	661	12 Q8QM64	Q8qm64 ectocarpus
27	66	13.9	707	11 Q9ESY6	Q9esy6 rattus norv
28	65	13.7	341	5 Q8IAX7	Q8iax7 plasmodium
29	65	13.7	416	5 Q18094	Q18094 caenorhabdi
30	64.5	13.6	174	2 Q8L344	Q8l344 vibrio chol
31	64.5	13.6	701	5 Q8IK93	Q8ik93 plasmodium
32	64	13.5	416	5 Q9N507	Q9n507 caenorhabdi
33	64	13.5	737	8 Q32744	Q32744 osteospermu
34	64	13.5	1293	10 Q9M8Z0	Q9m8z0 arabidopsis
35	63.5	13.4	340	11 Q9CXD6	Q9cxd6 mus musculu
36	63.5	13.4	505	5 Q8IDX5	Q8idx5 plasmodium
37	63.5	13.4	562	16 Q9KYG2	Q9kyg2 vibrio chol
38	63.5	13.4	612	16 Q8D3J0	Q8d3j0 wiggleswort
39	63.5	13.4	781	16 Q8EWN0	Q8ewn0 mycoplasma
40	63.5	13.4	1126	16 Q9NS6	Q9ns6 rhizobium l
41	63.5	13.4	1394	2 Q9AIU3	Q9aiu3 anaplasma p
42	63	13.3	290	9 Q8HA77	Q8ha77 bacterioph
43	63	13.3	294	16 Q92B18	Q92b18 listeria in
44	63	13.3	299	17 Q980R7	Q980r7 sulfolobus
45	63	13.3	334	5 Q8ILY8	Q8ily8 plasmodium

#### ALIGNMENTS

#### RESULT 1

Q9BRW0 PRELIMINARY; PRT; 240 AA.  
AC Q9BRW0;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Similar to CD58 antigen, (Lymphocyte function-associated antigen 3).  
DE Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005930; AAH05930.1; -  
DR HSP; PI9256; ICCZ.  
DR InterPro: IPR003599; Ig.  
DR SMART; SM00409; IG; 1.  
SQ SEQUENCE 240 AA; 27044 MW; 431E44EFDDF80E2 CRC64;

Query Match 100.0%; Score 475; DB 4; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.8e-46;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MVAGSDAGRALGVLSVWCLLHCFGFCISCFQQIYGVYGVNTHFVPSNPLKYLKKOK 60  
DB 1 MVAGSDAGRALGVLSVWCLLHCFGFCISCFQQIYGVYGVNTHFVPSNPLKYLKKOK 60  
QY 61 DKVAELNSEFRATSSFKNRNYLDTVSGSLTI 92  
DB 61 DKVAELNSEFRATSSFKNRNYLDTVSGSLTI 92

#### RESULT 2

Q14748



```

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE LFA-3(delta D2) precursor.
OS Ovis sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Etheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9939;
RN [1]
RP SEQUENCE FROM N.A.
RA Kakutani T.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D28583; BAA05921.1; -.
DR HSSP; P19256; ICCZ.
DR InterPro; IPR003599; IG.
DR SMART; SM00409; IG; 1.
KW Signal.
FT SIGNAL
FT FT
SQ SEQUENCE 159 AA; 17645 MW; 592A97046ECD985E CRC64;
Query Match 54.0%; Score 256.5; DB 6; Length 159;
Best Local Similarity 58.1%; Pred. No. 1.5e-21;
Matches 54; Conservative 9; Mismatches 29; Indels 1; Gaps 1;
QY 1 MVAGSDAGRGVLGVSVCLLHCFGFISCFQQIYGVVYGNVFHFVPSNVPLEKVLKKOK 60
Db | ||| |||| | : ||| |||| | : |||| | : | : ||| |
1 MAAGSAPGALCALGLVCLFLKLDFISCVSDIVGAMNGNVFYVESQPFFTEIMWKKGK 60
QY 61 DKVAEL-ENSEFRAFSPFKRNYLDTVSGSLTI 92
| | | : : | | | ||||| : | | | : | |
Db 61 DKVVEWDTSGLEAFQSFKNRVHLDIVSGNLTI 93
RESULT 5
Q28754 PRELIMINARY; PRT; 227 AA.
ID Q28754 AC Q28754;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE LFA-3(delta TM) precursor.
OS Ovis sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Etheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9939;
RN [1]
RP SEQUENCE FROM N.A.
RA Kakutani T.;
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; D28585; BAA05921.1; -.
DR HSSP; P19256; ICCZ.
DR InterPro; IPR003599; IG.
DR SMART; SM00409; IG; 1.
KW Signal.
FT SIGNAL
FT FT
SQ SEQUENCE 227 AA; 24760 MW; 1521A800B49E9B55 CRC64;
Query Match 54.0%; Score 256.5; DB 6; Length 227;
Best Local Similarity 58.1%; Pred. No. 2.3e-21;
Matches 54; Conservative 9; Mismatches 29; Indels 1; Gaps 1;
QY 1 MVAGSDAGRGVLGVSVCLLHCFGFISCFQQIYGVVYGNVFHFVPSNVPLEKVLKKOK 60
Db | ||| |||| | : ||| |||| | : |||| | : | : ||| |
1 MAAGSAPGALCALGLVCLFLKLDFISCVSDIVGAMNGNVFYVESQPFFTEIMWKKGK 60
QY 61 DKVAEL-ENSEFRAFSPFKRNYLDTVSGSLTI 92
| | | : : | | | ||||| : | | | : | |
Db 61 DKVVEWDTSGLEAFQSFKNRVHLDIVSGNLTI 93
RESULT 6

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:15:32 ; Search time 29 Seconds  
(without alignments)  
134.228 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_1\_92

Perfect score: 475

Sequence: 1 NVAGSDAGRALGVLSVVCLL.....AFSSFKRVYLDVSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	475	100.0	240	1	US-07-940-861-12
2	475	100.0	240	1	US-08-459-512-12
3	475	100.0	240	2	US-08-459-657-12
4	475	100.0	240	2	US-08-460-132-12
5	475	100.0	240	3	US-08-466-465-4
6	475	100.0	240	5	PCT-US92-02050-12
7	475	100.0	240	6	5185441-36
8	475	100.0	240	6	5223394-4
9	475	100.0	240	6	5223394-6
10	475	100.0	250	1	US-07-940-861-10
11	475	100.0	250	1	US-08-459-512-10
12	475	100.0	250	2	US-08-459-657-10
13	475	100.0	250	2	US-08-460-132-10
14	475	100.0	250	3	US-08-466-465-2
15	475	100.0	250	5	PCT-US92-02050-10
16	475	100.0	250	6	5223394-1
17	475	100.0	251	6	5185441-38
18	475	100.0	347	1	US-07-940-861-43
19	475	100.0	347	1	US-08-459-512-43
20	475	100.0	347	1	US-08-459-657-43
21	475	100.0	347	2	US-08-460-132-43
22	475	100.0	347	3	US-08-466-465-8
23	475	100.0	347	5	PCT-US92-02050-43
24	402	84.6	77	1	US-07-940-861-5
25	402	84.6	77	1	US-08-459-512-5
26	402	84.6	77	2	US-08-459-657-5
27	402	84.6	77	2	US-08-460-132-5

28 402 84.6 77 5 PCT-US92-02050-5  
29 328 69.1 134 1 US-08-328-152A-13  
30 328 69.1 222 1 US-08-328-152A-8  
31 260 54.7 50 1 US-07-940-861-2  
32 260 54.7 50 1 US-08-459-512-2  
33 260 54.7 50 2 US-08-459-657-2  
34 260 54.7 50 2 US-08-460-132-2  
35 260 54.7 50 5 PCT-US92-02050-2  
36 200 42.1 38 6 5185441-1  
37 186.5 39.3 131 1 US-08-328-152A-1  
38 186.5 39.3 199 1 US-08-328-152A-36  
39 186.5 39.3 225 1 US-08-328-152A-31  
40 74.5 15.7 194 3 US-08-630-172-14  
41 74.5 15.7 194 3 US-09-375-419-14  
42 73 15.4 16 1 US-07-940-861-7  
43 73 15.4 16 1 US-08-459-512-7  
44 73 15.4 16 2 US-08-459-657-7  
45 73 15.4 16 2 US-08-460-132-7

#### ALIGNMENTS

RESULT 1  
US-07-940-861-12  
; Sequence 12, Application US/07940861  
; Patent No. 5547853  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALINER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/940,861  
; FILING DATE: 21-OCT-1992  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

Sequence 5, Appli  
Sequence 13, Appli  
Sequence 8, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Patent No. 5185441  
Sequence 1, Appli  
Sequence 36, Appli  
Sequence 31, Appli  
Sequence 14, Appli  
Sequence 14, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli

## MOLECULE TYPE: protein

JS-07-940-861-12

Query Match 100.0%; Score 475; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.1e-51;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2Y 1 MVAGSDAGRALGVLVSVVCLLHCFGFCFSCFSQQIYGVVGNVTHFVPSNVLKVLKWKOK 60  
|||||  
Db 1 MVAGSDAGRALGVLVSVVCLLHCFGFCFSCFSQQIYGVVGNVTHFVPSNVLKVLKWKOK 60  
|||||  
2Y 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||  
Db 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

## RESULT 2

US-08-459-512-12  
; Sequence 12, Application US/08459512  
; Patent No. 5728677

GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Neave  
STREET: 875 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10022-6250

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,512  
FILING DATE: 02-JUN-1995

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02050

FILING DATE: 12-MAR-1992  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,512  
FILING DATE: 02-JUN-1995

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02050

FILING DATE: 12-MAR-1992  
APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,967  
FILING DATE: 07-OCT-1991

ATTORNEY/AGENT INFORMATION:  
NAME: HALEY, James F., Jr.  
REGISTRATION NUMBER: 27,794

REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)715-0600

TELEFAX: (212)715-0673  
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids

TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-459-512-12

Query Match 100.0%; Score 475; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.1e-51;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MVAGSDAGRALGVLVSVVCLLHCFGFCFSCFSQQIYGVVGNVTHFVPSNVLKVLKWKOK 60  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

Db 1 MVAGSDAGRALGVLVSVVCLLHCFGFCFSCFSQQIYGVVGNVTHFVPSNVLKVLKWKOK 60  
|||||

QY 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

Db 61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

61 DKVAELENSEFRASFSSKRNRYLDTVSGSLTI 92  
|||||

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:17:27 ; Search time 67 seconds  
(without alignments)  
217.248 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_1\_92

Perfect score: 475

Sequence: 1 MVAGSDAGRALGVLSVCLL.....AFSFKRNRYLDIVSGSLTI 92

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 587654 segs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	475	100.0	240	9	US-09-796-033-4
2	475	100.0	240	9	US-09-730-465-4
3	475	100.0	250	9	US-09-796-033-2
4	475	100.0	250	9	US-09-730-465-2
5	475	100.0	250	12	US-10-436-523-61
6	475	100.0	250	15	US-10-207-655-113
7	475	100.0	347	9	US-09-796-033-8
8	475	100.0	347	9	US-09-730-465-8
9	475	100.0	347	14	US-10-091-236-17
10	475	100.0	347	15	US-10-091-313-7
11	475	100.0	347	15	US-10-091-268-7
12	74.5	15.7	96	9	US-09-864-761-47502
13	74.5	15.7	243	11	US-09-860-836B-13
14	74.5	15.7	243	12	US-10-436-523-62
15	74.5	15.7	243	15	US-10-207-655-111

16	71	14.9	287	12	US-10-032-214-237	Sequence 237, App
17	70	14.7	288	12	US-10-032-214-72	Sequence 72, Appl
18	70	14.7	288	12	US-10-032-214-82	Sequence 82, Appl
19	70	14.7	288	12	US-10-032-214-83	Sequence 83, Appl
20	70	14.7	288	12	US-10-032-214-93	Sequence 93, Appl
21	70	14.7	288	12	US-10-032-214-221	Sequence 227, App
22	70	14.7	288	12	US-10-032-214-246	Sequence 246, App
23	70	14.7	288	12	US-10-032-214-251	Sequence 251, App
24	70	14.7	288	12	US-10-032-214-252	Sequence 252, App
25	70	14.7	288	12	US-10-032-214-279	Sequence 279, App
26	68	14.3	287	12	US-10-032-214-74	Sequence 74, Appl
27	68	14.3	288	12	US-10-032-214-75	Sequence 75, Appl
28	68	14.3	288	12	US-10-032-214-77	Sequence 77, Appl
29	68	14.3	288	12	US-10-032-214-78	Sequence 78, Appl
30	68	14.3	288	12	US-10-032-214-87	Sequence 87, Appl
31	68	14.3	288	12	US-10-032-214-91	Sequence 91, Appl
32	68	14.3	288	12	US-10-032-214-92	Sequence 92, Appl
33	68	14.3	288	12	US-10-032-214-234	Sequence 234, App
34	68	14.3	288	12	US-10-032-214-247	Sequence 247, App
35	68	14.3	288	12	US-10-032-214-286	Sequence 286, App
36	67.5	14.2	240	11	US-09-860-836B-15	Sequence 15, Appl
37	67	14.1	288	12	US-10-032-214-85	Sequence 85, Appl
38	67	14.1	288	12	US-10-032-214-88	Sequence 88, Appl
39	67	14.1	288	12	US-10-032-214-289	Sequence 289, App
40	66	13.9	288	12	US-10-032-214-287	Sequence 287, App
41	66	13.9	707	12	US-09-354-342-43	Sequence 43, Appl
42	65	13.7	288	12	US-10-032-214-288	Sequence 288, App
43	64	13.5	288	12	US-10-032-214-235	Sequence 235, App
44	63.5	13.4	351	11	US-09-829-275-2	Sequence 2, Appl
45	63	13.3	287	12	US-10-032-214-243	Sequence 243, App

#### ALIGNMENTS

#### RESULT 1

US-09-796-033-4  
; Sequence 4, Application US/09796033  
; Patent No. US20020009446A1  
; GENERAL INFORMATION:  
; APPLICANT: Magilavay, Daniel  
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR  
; FILE OF INVENTION: T-CELLS AND COMPOSITIONS  
; FILE REFERENCE: 10274-044001  
; CURRENT APPLICATION NUMBER: US/09/796.033  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/US99/20026  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: US 60/098,456  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(28)  
US-09-796-033-4

Query Match 100.0%; Score 475; DB 9; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.7e-49;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFCISFCSQIYGVYGVNTHVPSNVLKVKKQK 60

DB 1 MVAGSDAGRALGVLSVCLLHCFGFCISFCSQIYGVYGVNTHVPSNVLKVKKQK 60

QY 61 DKVAELENSEFRAFSSFKRNRYLDIVSGSLTI 92

DB 61 DKVAELENSEFRAFSSFKRNRYLDIVSGSLTI 92

## RESULT 2

S-09-730-465-4

Sequence 4, Application US/09730465

Patent No. US20020009449A1

GENERAL INFORMATION:

APPLICANT: Wallner, Barbara P.

Cooper, Kevin D.

TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen  
Presenting Cell Driven Skin Conditions Using

Inhibitors of the CD2/LFA-3 Interaction

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE &amp; COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/730,465

FILING DATE: 05-Dec-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/08755

FILING DATE: 06-OCT-1992

APPLICATION NUMBER: US 07/862,022

FILING DATE: 12-APR-1992

APPLICATION NUMBER: US 07/770,969

FILING DATE: 07-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis (PLM)

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-111CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-730-465-4

Query Match 100.0%; Score 475; DB 9; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.7e-49;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQIYGVVGNVTFHVPSPNPLKEVLKKQK 60

Db 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQIYGVVGNVTFHVPSPNPLKEVLKKQK 60

QY 61 DKVAELENSEFRAPSSFKNRVYLDTVSGSLTI 92

Db 61 DKVAELENSEFRAPSSFKNRVYLDTVSGSLTI 92

## RESULT 3

US-09-796-033-2

Sequence 2, Application US/09796033

Patent No. US20020009446A1

GENERAL INFORMATION:

APPLICANT: Magilavy, Daniel

TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR

T-CELLS AND COMPOSITIONS

FILE REFERENCE: 10274-044001

CURRENT APPLICATION NUMBER: US/09/796,033

; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/US99/20026  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: US 60/098,456  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 250  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(28)  
; US-09-796-033-2

Query Match 100.0%; Score 475; DB 9; Length 250;  
Best Local Similarity 100.0%; Pred. No. 1.8e-49;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQIYGVVGNVTFHVPSPNPLKEVLKKQK 60

Db 1 MVAGSDAGRALGVLSVCLLHCFGFTSCFSQIYGVVGNVTFHVPSPNPLKEVLKKQK 60

QY 61 DKVAELENSEFRAPSSFKNRVYLDTVSGSLTI 92

Db 61 DKVAELENSEFRAPSSFKNRVYLDTVSGSLTI 92

## RESULT 4

US-09-730-465-2

Sequence 2, Application US/09730465

Patent No. US20020009449A1

GENERAL INFORMATION:

APPLICANT: Wallner, Barbara P.

Cooper, Kevin D.

TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen  
Presenting Cell Driven Skin Conditions Using

Inhibitors of the CD2/LFA-3 Interaction

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE &amp; COCKFIELD

STREET: 60 State Street, Suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/730,465

FILING DATE: 05-Dec-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/08755

FILING DATE: 06-OCT-1992

APPLICATION NUMBER: US 07/862,022

FILING DATE: 12-APR-1992

APPLICATION NUMBER: US 07/770,969

FILING DATE: 07-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: Myers, Louis (PLM)

REGISTRATION NUMBER: 35,965

REFERENCE/DOCKET NUMBER: BGP-111CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 250 amino acids

TYPE: amino acid

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

3M protein - protein search, using sw model

Run on: October 7, 2003, 16:34:43 ; Search time 17.9412 Seconds  
(without alignments)  
539.670 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_20\_80

Perfect score: 327  
Sequence: 1 LHCFGFCFSQQYGVVG.....DKVAELENSEFRAPSSFKNR 61

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	112	23 ABB81992	Human transmembran
2	327	100.0	120	13 AAR27157	LFA-3 CD2 binding
3	327	100.0	128	17 AAW04361	Human cell adhesio
4	327	100.0	133	17 AAW04362	Human cell adhesio
5	327	100.0	133	17 AAW04363	Human cell adhesio
6	327	100.0	237	12 AAR14182	Human LFA-3(CD58)
7	327	100.0	237	13 AAR20804	Human LFA-3 antige
8	327	100.0	237	17 AAR91432	Human LFA-3 antige
9	327	100.0	237	19 AAW80441	Human LFA-3 antige

10	327	100.0	237	20 AAW86189	Human LFA-3 antige
11	327	100.0	237	21 AAY96127	Human cell surface
12	327	100.0	237	22 AAU02436	Human lymphocyte f
13	327	100.0	240	11 AAR05572	PI-linked LFA-3 en
14	327	100.0	240	11 AAR07604	Lymphocyte functio
15	327	100.0	240	13 AAR27162	PI-linked LFA-3 us
16	327	100.0	240	14 AAR34372	Sequence encoded b
17	327	100.0	240	21 AAY83134	PI-linked human tr
18	327	100.0	240	22 AAB61158	Human PI-linked tr
19	327	100.0	240	23 AAU76226	PI-linked lymphocy
20	327	100.0	250	9 AAP91507	Human LFA-3 expres
21	327	100.0	250	13 AAR27161	transmembrane LFA-
22	327	100.0	250	14 AAR34371	Sequence encoded b
23	327	100.0	250	15 AAR64271	Human LFA-3. Homo
24	327	100.0	250	17 AAW04370	Human cell adhesio
25	327	100.0	250	21 AAY83133	Human transmembran
26	327	100.0	250	22 AAB61157	Human transmembran
27	327	100.0	250	23 AAU76225	Lymphocyte functio
28	327	100.0	256	24 AA016017	Human transmembran
29	327	100.0	280	22 AAG18595	Murine CD2-binding
30	327	100.0	347	13 AAR27163	Novel human disigno
31	327	100.0	347	21 AAY83136	CD2 binding LFA-3-
32	327	100.0	347	22 AAB61160	Human transmembran
33	327	100.0	347	23 AAG70766	Human LFA3TIP fusi
34	327	100.0	347	23 ABB81994	Human lymphocyte f
35	327	100.0	347	23 ABB81994	Human LFA3TIP poly
36	327	100.0	347	24 APU76228	LFA3/IgG fusion pr
37	327	100.0	351	23 ABB81995	Murine LFA-3. Mus
38	324	99.1	240	14 AAR34222	Human LFA-3/IgG fu
39	324	99.1	250	14 AAR34221	Amino acid sequenc
40	321	98.2	237	18 AAW16687	Human CD58 GPI. H
41	319	97.6	347	14 AAR34224	Amino acid sequenc
42	317	96.9	250	13 AAR28365	Human LFA-3 protei
43	271	82.9	71	13 AAR27158	LFA-3 CD2 binding
44	271	82.9	134	13 AAR28364	Human LFA-3 D2 reg
45	271	82.9	135	13 AAR28369	Human LFA-3-delta-

#### ALIGNMENTS

##### RESULT 1

ABB81992  
ID ABB81992 standard; Protein; 112 AA.

XX AC ABB81992;

XX DT 25-NOV-2002 (first entry)

XX DE Human transmembrane LFA-3 partial polypeptide.

XX KW Human; immunosuppressive; antithyroid; dermatological; antiinflammatory;  
antipsoriatic; antirheumatic; antiarthritic; neuroprotective; vaccine;  
osteopathic; antidiabetic; gene therapy; CD2; LFA-3; interferon; IFN;  
CD25; CD8; CD45.

XX OS Homo sapiens.

XX PH Location/Qualifiers

XX FT Key 1..28

XX FT Peptide /note= "signal peptide"

XX FT Protein 29..84

XX FT /note= "partial mature protein fragment; the sequence

after position 84 is not indicated in the

specification; the pages containing the complete

sequence are not provided"

XX FT WO200260480-A1.

XX FT 08-AUG-2002.

XX FT 25-JAN-2002; 2002WO-US02314.

3 01-FEB-2003; 2001US-265964P.  
 4 (BIOJ ) BIOGEN INC.  
 5 Vaishnav AK, Cooper KD, Shrager D, McCormick TS;  
 6 N-PSDB; ABQ79635.  
 7  
 8 Treating or preventing, by inhibiting CD2/LFA-3 interaction, an  
 9 epidermal or dermal disorder, inflammatory and/or autoimmune disorder  
 10 with aberrant T cell activity or proliferation.  
 11  
 12 Disclosure; Page 58; 68pp; English.  
 13  
 14 The invention relates to treating or preventing an epidermal or dermal  
 15 disorder, an inflammatory disorder and/or an autoimmune disorder in a  
 16 subject with aberrant T cell activity or proliferation. The method  
 17 involves administering an inhibitor of the CD2/LFA-3 interaction, in  
 18 combination with an auxiliary agent, thereby treating or preventing the  
 19 epidermal or dermal disorder, the inflammatory disorder or the autoimmune  
 20 disorder. The methods and compositions of the invention are useful for  
 21 preventing or treating skin disorders characterized by increased T cell  
 22 activation and abnormal antigen presentation in the dermis and epidermis,  
 23 such as psoriasis, UV damage, atopic dermatitis, cutaneous T cell  
 24 lymphoma, contact and atopic dermatitis, lichen planus, alopecia,  
 25 vitiligo, urticaria. They may also be used in chronic inflammatory and  
 26 autoimmune disorders such as diabetes mellitus, arthritis, rheumatoid  
 27 arthritis, juvenile rheumatoid arthritis, osteoarthritis, psoriatic  
 28 arthritis, multiple sclerosis, encephalomyelitis, myasthenia gravis,  
 29 systemic lupus erythematosus, autoimmune thyroiditis. The present  
 30 sequence represents the human transmembrane LFA-3 partial amino acid  
 31 sequence.  
 32  
 33 Sequence 112 AA;  
 34  
 35 Query Match 100.0%; Score 327; DB 23; Length 112;  
 36 Best Local Similarity 100.0%; Pred. No. 2.2e-34;  
 37 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 38  
 39 1 LHCFGFISCFISQIIGVYGVNTHFVPSNPLKEVLWKQKDKVAELENSEFRASFKN 60  
 40 |  
 41 20 LHCFGFISCFISQIIGVYGVNTHFVPSNPLKEVLWKQKDKVAELENSEFRASFKN 79  
 42  
 43 61 R 61  
 44 80 R 80  
 45  
 46 RESULT 2  
 47 AAR27157  
 48 ID AAR27157 standard; protein; 120 AA.  
 49 XX  
 50 AC AAR27157;  
 51 XX  
 52 XX  
 53 25-MAR-2003 (updated)  
 54 DT 20-MAY-1998 (first entry)  
 55 XX  
 56 XX  
 57 DE LFA-3 CD2 binding domain polypeptide #2.  
 58 XX  
 59 lymphocyte associated antigen-3; T-lymphocyte accessory molecule.  
 60 XX  
 61 OS Synthetic.  
 62 XX  
 63 Key Location/Qualifiers  
 64 FT region /note= "1-77 of these residues may be deleted"  
 65 FT region 89..120  
 66 FT /note= "1-32 of these residues may be deleted"  
 67 XX  
 68 XX EP503648-A1.  
 69 XX  
 70 16-SEP-1992.  
 71 XX

XX 12-MAR-1992; 92EP-0104320.  
 XX  
 XX 12-MAR-1991; 91US-0667971.  
 PR 07-OCT-1991; 91US-0770967.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 XX  
 PI Miller GT, Rosa MD, Wallner BP;  
 XX  
 DR WPI; 1992-309760/38.  
 XX  
 XX CD2-binding domain of lymphocyte function associated antigen-3  
 PT and DNA - for diagnosing and treating inflammation and  
 PT autoimmune diseases, e.g. systemic lupus erythematosus and  
 PT rheumatoid arthritis  
 XX  
 PS Claim 1; Page 52-53; 85pp; English.  
 XX  
 CC This polypeptide can bind to CD2. The N-terminal Met is opt.  
 CC absent. The N-terminal and C-terminal portions can be deleted.  
 CC the polypeptide and its functional deletion mutants may be used to  
 CC treat acute and chronic inflammation, autoimmune disease and for  
 CC immunomodulation. The polypeptides can also form the N-terminal  
 CC part of a fusion protein. The polypeptides and fusion proteins may  
 CC also be used to inhibit T-cell activation and the proliferation of  
 CC peripheral blood lymphocytes. Multimeric proteins can be formed from  
 CC the polypeptides and/or fusion proteins. The multimers have enhanced  
 CC affinity for CD2. See also AAR27157-9. field.)  
 CC (updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX Sequence 120 AA;  
 30  
 31 Query Match 100.0%; Score 327; DB 13; Length 120;  
 32 Best Local Similarity 100.0%; Pred. No. 2.4e-34;  
 33 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 34  
 35 QY 1 LHCFGFISCFISQIIGVYGVNTHFVPSNPLKEVLWKQKDKVAELENSEFRASFKN 60  
 36 |  
 37 20 LHCFGFISCFISQIIGVYGVNTHFVPSNPLKEVLWKQKDKVAELENSEFRASFKN 79  
 38  
 39 QY 61 R 61  
 40 80 R 80  
 41  
 42 RESULT 3  
 43 AAW04361  
 44 ID AAW04361 standard; Protein; 128 AA.  
 45 XX  
 46 XX  
 47 AC AAW04361;  
 48 XX  
 49 XX 27-JUN-1997 (first entry)  
 50 XX  
 51 XX Human cell adhesion protein LFA-3.  
 52 XX  
 53 XX Human; T cell; MOLT-4; cell adhesion; LFA-3; immunosuppression;  
 54 XX cell line; immunosuppressant cell.  
 55 XX  
 56 OS Homo sapiens.  
 57 XX  
 58 XX W09633217-A1.  
 59 XX  
 60 XX 24-OCT-1996.  
 61 XX  
 62 XX 15-APR-1996; 96WO-JP01039.  
 63 PF  
 64 XX 27-DEC-1995; 95JP-0341959.  
 65 PR 19-APR-1995; 95JP-0094060.  
 66 PR 04-JUL-1995; 95JP-0169110.  
 67 XX  
 68 XX (KANF ) KANEGAFUCHI KAGAKU KOGYO KK.  
 69 PA  
 70 XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

XM protein - protein search, using sw model

run on: October 7, 2003, 16:45:48 ; Search time 7.47549 Seconds  
(without alignments)  
784.736 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_20\_80

Perfect score: 327

Sequence: 1 LHCFGFISCFQQIYGVYG.....DKVAELENSEFRASFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: PIR\_76:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	250	2 A28564	lymphocyte functio
2	67	20.5	546	2 T26568	hypothetical prote
3	63.5	19.4	562	2 C82355	hypothetical prote
4	62	19.0	416	2 T25036	hypothetical prote
5	60.5	18.5	275	2 E33655	signal peptidase-1
6	60.5	18.5	508	1 A30007	golichyl-diphospho
7	60.5	18.5	1822	2 S44849	K12H4.8 protein -
8	60.5	18.5	2924	2 T18378	variant-specific s
9	60	18.3	405	2 S41917	methionine adenosy
10	60	18.3	619	2 C96714	unknown protein t6
11	60	18.3	852	2 A00050	probable phosphoen
12	59.5	18.2	273	2 B4608	outer membrane pro
13	59	18.0	294	2 A01649	weakly phase relat
14	59	18.0	344	2 T00993	probable beta-1,3-
15	59	18.0	475	2 T12955	probable protein k
16	59	18.0	511	2 B81332	GMP synthase (glut
17	58.5	17.9	508	2 F71865	GMP synthetase - H
18	58.5	17.9	656	2 S34197	DNA-binding protei
19	58	17.7	237	2 C70157	hypothetical prote
20	58	17.7	299	2 E90162	conserved hypothe
21	58	17.7	679	2 S54299	transketolase (EC
22	58	17.7	719	2 D90083	probable helicase-
23	58	17.7	737	2 T13504	NADH2 dehydrogenas
24	57.5	17.6	401	2 E69501	sugar transporter
25	57.5	17.6	425	2 T24111	hypothetical prote
26	57.5	17.6	632	2 C81261	transketolase (EC
27	57.5	17.6	786	2 AG2375	WD-40 repeat-prote
28	57.5	17.6	1157	2 T40572	protein phosphatas
29	57	17.4	509	2 D86146	F22L4.11 protein -

30 56.5 17.3 299 2 D81871 RdcG protein NMA10  
31 56.5 17.3 299 2 F81149 recombination asso  
32 56.5 17.3 1656 2 S54520 probable membrane  
33 56 17.1 142 2 E71079 hypothetical prote  
34 56 17.1 265 2 G75067 hypothetical prote  
35 56 17.1 274 2 S76154 hypothetical prote  
36 56 17.1 378 2 G96630 omega-6 fatty acid  
37 56 17.1 382 2 T15042 probable enoyl-CoA  
38 56 17.1 601 2 T38200 spliceosome associ  
39 56 17.1 745 2 S57048 hypothetical prote  
40 56 17.1 752 2 H97191 uncharacterized pr  
41 56 17.1 918 2 C84720 hypothetical prote  
42 56 17.1 1107 1 S52517 myosin I heavy cha  
43 56 17.1 1363 2 T43220 insulin-like growt  
44 55.5 17.0 348 2 B64552 conserved hypothet  
45 55.5 17.0 351 2 D71955 hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A28564  
lymphocyte function-associated antigen 3, transmembrane splice form precursor - hum  
N:Alternate names: CD58 antigen; surface glycoprotein LFA-3  
C:Species: Homo sapiens (man)  
C>Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 02-Aug-2002  
C:Accession: A28564; S01269  
R:Wallner, B.P.; Frey, A.Z.; Tizard, R.; Mattaliano, R.J.; Hession, C.; Sanders, M.  
J. Exp. Med. 166, 923-932, 1987  
A:Title: Primary structure of lymphocyte function-associated antigen 3 (LFA-3). The  
A:Reference number: A28564; MUID:88009714; PMID:3309127  
A:Accession: A28564  
A:Molecule type: mRNA  
A:Residues: 1-250 <WALL>  
A:Cross-references: DB:Y00636; NID:g34346; PIDN:CAA68668.1; PID:g34347  
A:Experimental source: erythrocytes  
A:Note: the authors translated the codon TAT for residue 34 as Val  
R:Seed, B.  
Nature 329, 840-842, 1987  
A:Title: An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to  
A:Reference number: S01269; MUID:88039074; PMID:3313052  
A:Accession: S01269  
A:Molecule type: mRNA  
A:Residues: 1-235, 'VL' <SEE>  
A:Cross-references: EMBL:X06296; NID:g34349; PIDN:CAA29622.1; PID:g34350  
C:Comment: For an alternative splice form, see PIR:S01269.  
C:Comment: For an alternative splice form, see PIR:A28564.  
C:Genetics:  
A:Gene: GDB:CD58; LFA3  
A:Cross-references: GDB:120580; OMIM:153420  
A:Map position: lp13-lp13  
C:Superfamily: human B-cell antigen CD19; immunoglobulin homology  
C:Keywords: alternative splicing; glycoprotein; phosphatidylinositol linkage; surfa  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-250/Product: lymphocyte function-associated antigen 3 transmembrane splice for  
F:135-189/Domain: immunoglobulin homology <IMM>  
F:40,94,109,135,169,195/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 100.0%; Score 327; DB 2; Length 250;

Best Local Similarity 100.0%; Pred. No. 7.1e-32;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LHCFGFISCFQQIYGVYG...DKVAELENSEFRASFKNR 60

DB 20 LHCFGFISCFQQIYGVYG...DKVAELENSEFRASFKNR 79

QY 61 R 61

DB 80 R 80

##### RESULT 2



A:Molecule type: DNA  
A:Residues: 1-416 <EML>  
A:Cross-references: EMBL:Z81594; PTDN:CAB04746.1; GSPDB:GN00019; CESP:T20F10.4  
A:Experimental source: clone T20F10  
C:Genetics:  
A:Gene: CESP:T20F10.4  
A:Map position: 1

Query Match 19.0%; Score 62; DB 2; Length 416;  
Best Local Similarity 35.7%; Pred. No. 13;  
Matches 10; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 2 HCFGPISCSQIYGVYGVNVTFFVPSN 29  
| | | | | : : : | | : | |  
Db 165 HIFGVEVCYNKIYGRKHNLSLYRSPN 192

RESULT 5  
E83655  
signal peptidase-like protein BH0045 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: E83655  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.;  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: E83655  
A:Status: preliminary  
A:Superfamily: probable signal peptidase II yaar

A:Molecule type: DNA  
A:Residues: 1-275 <STO>  
A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PTDN:BA03764.1; GSPDB  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0045  
C:Superfamily: probable signal peptidase II yaar

Query Match 18.5%; Score 60.5; DB 2; Length 275;  
Best Local Similarity 44.2%; Pred. No. 12;  
Matches 19; Conservative 5; Mismatches 12; Indels 7; Gaps 3;

QY 16 GYVYGVNVTFFH--VPSN--VPLKEVL-WKKQDKVAELENSE 51  
| | | | | | | | | | : | | | | : | | | | | | | |  
Db 38 GVEYGVNVIEKKTGVDNDVDVLPKQVIVVATEKDKLAVQENKE 80

RESULT 6  
A30007  
dolichyl-diphosphooligosaccharide-protein glycotransferase (EC 2.4.1.119) glycosylat  
C:Species: Gallus gallus (chicken)  
C:Date: 15-Dec-1988 #sequence\_revision 26-May-1994 #text\_change 03-Jun-2002  
C:Accession: A30007  
R:Geetha-Habib, M.; Noiva, R.; Kaplan, H.A.; Lennarz, W.J.  
Cell 54, 1053-1060, 1988  
A:Title: Glycosylation site binding protein, a component of oligosaccharyl transfer  
A:Reference number: A30007; MUID:88327849; PMID:2458190  
A:Accession: A30007  
A:Molecule type: mRNA  
A:Residues: 1-508 <GEE>  
C:Cross-references: EMBL:M25994  
C:Superfamily: protein disulfide-isomerase; thioredoxin homology  
C:Keywords: duplication; endoplasmic reticulum; glycosyltransferase; hexosyltransfer  
F:1-17/Domain: signal sequence #status predicted <SIG>  
F:18-508/Product: dolichyl-diphosphooligosaccharide-protein glycotransferase glycosy  
F:31-118/Domain: thioredoxin homology <TXL>  
F:376-461/Domain: thioredoxin homology <TX2>  
F:505-508/Region: endoplasmic reticulum retention signal  
F:52-55,398-401/Disulfide bonds: redox-active #status predicted  
F:312-343/Disulfide bonds: #status predicted

Query Match 18.5%; Score 60.5; DB 1; Length 508;  
Best Local Similarity 43.9%; Pred. No. 24;  
Matches 18; Conservative 3; Mismatches 19; Indels 1; Gaps 1;

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

M protein - protein search, using sw model

run on: October 7, 2003, 16:35:28 ; Search time 4.03677 Seconds

(without alignments)  
710.626 Million cell updates/sec

File: US-09-730-465-2\_COPY\_20\_80

Perfect score: 327

Sequence: 1 LHCFGTFSCFSQRIYGVYG.....DKVAELENSEFRAGSFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	327	100.0	250	1 LPA3_HUMAN	P19256 homo sapien
2	61.5	18.8	379	1 ASPN_HUMAN	Q9bxn1 homo sapien
3	60.5	18.5	373	1 ASPN_MOUSE	Q99mq4 mus musculu
4	60.5	18.5	508	1 GSBP_CHICK	P12244 gallus gall
5	60	18.3	408	1 METK_DROME	P40320 drosophila
6	59	18.0	438	1 ARLY_CLOTE	P59616 clostridium
7	59	18.0	511	1 GUAA_CAMJE	Q9pn49 campylobact
8	58.5	17.9	508	1 GUAA_HELPJ	Q9zkg4 helicobacte
9	58	17.7	679	1 TETA_CRAPL	Q42675 craterostig
10	57.5	17.6	552	1 YHLL_SCHPO	Q9c0v8 schizosacch
11	56.5	17.3	299	1 RDGC_NEIGO	Q87408 neisseria g
12	56.5	17.3	299	1 RDGC_NEIMA	Q9jv02 neisseria m
13	56.5	17.3	299	1 RDGC_NEIMA	Q9jv02 neisseria m
14	56.5	17.3	974	1 ATXB_LEIDO	P12532 leishmania
15	56.5	17.3	1656	1 ATC8_YEAST	Q12674 saccharomyc
16	56	17.1	745	1 Y100_YEAST	P47101 saccharomyc
17	56	17.1	1107	1 MYLE_RAT	Q63356 rattus norv
18	56	17.1	1363	1 ILPR_BRALA	Q02466 branchiost
19	55.5	17.0	348	1 Y258_HELPJ	P56136 helicobacte
20	55.5	17.0	350	1 Y258_HELPJ	Q9zmh8 helicobacte
21	55.5	17.0	974	1 C15_YEAST	P27636 saccharomyc
22	55	16.8	651	1 Y942_METJA	Q58352 methanococ
23	55	16.8	844	1 MCEL_VARY	P33057 variola vir
24	55	16.8	950	1 Q9PL_HUMAN	Q9bxw6 homo sapien
25	54.5	16.7	195	1 MOBA_YERPE	Q8zjs4 yersinia pe
26	54.5	16.7	544	1 MLH1_HORVU	Q49873 hordeum vul
27	54.5	16.7	898	1 TOP1_SYNP3	P73810 synecocyst
28	54.5	16.7	974	1 ARXA_LEIDO	P11718 leishmania
29	54	16.5	833	1 PIAT_ECOLI	P32670 escherichia
30	54	16.5	1010	1 CONT_CHICK	P44781 gallus gall
31	54	16.5	1109	1 MYLE_HUMAN	Q12965 homo sapien
32	53.5	16.4	210	1 RL13_MOUSE	P47963 mus musculu
33	53.5	16.4	304	1 PAGO_SALTY	O30646 salmonella

34	53.5	16.4	560	1 ASNS_RAT	P49088 rattus norv
35	53.5	16.4	681	1 TRE_MANSE	P22297 manduca sex
36	53.5	16.4	1118	1 YIPL_YEAST	P40456 saccharomyc
37	53.5	16.4	1178	1 RPOB_MYCLE	P30760 mycobacteri
38	53.5	16.4	1178	1 RPOB_MYCLE	P47766 mycobacteri
39	53	16.2	212	1 SODF_STRCO	O51917 streptomyc
40	53	16.2	329	1 YX99_BACHD	Q9k7g2 bacillus ha
41	53	16.2	340	1 VB17_VARY	P33878 variola vir
42	53	16.2	368	1 PGS1_XENLA	Q9ib75 xenopus lae
43	53	16.2	474	1 CBLC_HUMAN	Q9ulv8 homo sapien
44	53	16.2	490	1 YAM7_SCHPO	Q10062 schizosacch
45	53	16.2	526	1 SI7A_MOUSE	Q9qz39 mus musculu

## ALIGNMENTS

RESULT 1	LEPA3_HUMAN	STANDARD;	PRT;	250 AA.
ID	LPA3_HUMAN	STANDARD;	PRT;	250 AA.
AC	P19256; Q96KI9;			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Lymphocyte function-associated antigen 3 precursor (Ag3) (Antigen			
DE	CD58) (Surface glycoprotein LFA-3).			
GN	CD58 OR LFA3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.			
RC	TISSUE=Erythrocyte;			
RX	MEDLINE=88009714; PubMed=3309127;			
RA	Wallner B.P., Frey A.Z., Tizard R., Mattaliano R.J., Hession C.,			
RA	Sanders M.E., Dustin M.L., Springer T.A.;			
RT	"Primary structure of the lymphocyte function-associated antigen 3 (LFA-			
RT	3). The ligand of the T lymphocyte CD2 glycoprotein.";			
RL	J. Exp. Med. 166:923-932(1987).			
[2]				
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RX	MEDLINE=88039074; PubMed=3313052;			
RA	Seed B.;			
RT	"An LFA-3 cDNA encodes a phospholipid-linked membrane protein			
RL	homologous to its receptor CD2.";			
Nature	329:840-842(1987).			
[3]				
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).			
MEDLINE=96169008;	PubMed=9510189;			
RA	Wallich R., Brenner C., Brand Y., Roux M., Reister M., Meuer S.;			
RT	"Gene structure, promoter characterization, and basis for alternative			
RNA	splicing of the human CD58 gene.";			
J. Immunol.	160:2862-2871(1998).			
[4]				
RP	X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 29-122.			
RX	MEDLINE=99218273; PubMed=10200255;			
RA	Ikemizu S., Sparks L.M., van der Merwe P.A., Harlos K., Stuart D.I.,			
RA	Jones E.Y., Davis S.J.;			
RT	"Crystal structure of the CD2-binding domain of CD58 (lymphocyte			
function-associated antigen 3) at 1.8-A resolution.";				
Proc. Natl. Acad. Sci. U.S.A.	96:4289-4294(1999).			
CC	-!- FUNCTION: LIGAND OF THE T LYMPHOCYTE CD2 GLYCOPROTEIN. THIS			
INTERACTION IS IMPORTANT IN MEDIATING THYMOCYTE INTERACTIONS WITH				
THYMIC EPITHELIAL CELLS, ANTIGEN-INDEPENDENT AND -DEPENDENT				
INTERACTIONS OF T LYMPHOCYTES WITH TARGET CELLS AND ANTIGEN-				
PRESENTING CELLS AND THE T LYMPHOCYTE ROSETTING WITH ERYTHROCYTES.				
IN ADDITION, THE LFA-3/CD2 INTERACTION MAY PRIME RESPONSE BY BOTH				
THE CD2+ AND LFA-3+ CELLS.				
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1) and			
attached to the membrane by a GPI-anchor (isoform 2).				
CC	-!- ALTERNATIVE PRODUCTS:			
Event=Alternative splicing; Named isoforms=2;				

```
Name=Long;
IsoId=PI9256-1; Sequence=Displayed;
Name=Short;
IsoId=PI9256-2; Sequence=VSP_002522, VSP_002523;
-!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
-!- DATABASE: NAME=PRO; NOTE=CD guide CD58 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd58.htm".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Y00636; CAA68668.1; -
EMBL; X06296; CAA29622.1; -
EMBL; Y14780; CAA75083.1; -
EMBL; Y14781; CAA75083.1; JOINED.
EMBL; Y14782; CAA75083.1; JOINED.
EMBL; Y14783; CAA75083.1; JOINED.
EMBL; Y14784; CAA75083.1; JOINED.
EMBL; Y14781; CAA75084.1; -
EMBL; Y14782; CAA75084.1; JOINED.
EMBL; Y14783; CAA75084.1; JOINED.
EMBL; Y14784; CAA75084.1; JOINED.
EMBL; Y14785; CAA75084.1; JOINED.
PIR; A28564; A28564.
PDB; 1CCZ; 05-APR-99.
PDB; 1C15; 22-JUN-99.
Genew; HGNC:1688; CD58.
MIM; 153420; -.
GO; GO:0005887; C:integral to plasma membrane; NAS.
GO; GO:0005515; P:protein binding activity; IPI.
GO; GO:0016337; P:cell-cell adhesion; NAS.
InterPro; IPR003599; IG.
SMART; SM00409; IG; 1.
PROSITE; PS00835; IG LIKE; FALSE_NEG.
Immunoglobulin domain; Glycoprotein; Transmembrane; Signal; GPI-anchor;
Alternative splicing; 3D-structure.
SIGNAL 1 28
CHAIN 29 250 LYMFOCYTE FUNCTION-ASSOCIATED ANTIGEN 3.
DOMAIN 29 215 EXTRACELLULAR (POTENTIAL).
TRANSMEM 216 238 POTENTIAL.
DOMAIN 239 250 CYTOPLASMIC (POTENTIAL).
DOMAIN 135 194 IG-LIKE C2-TYPE.
DISULFID 142 187 BY SIMILARITY.
CARBOHYD 40 40 N-LINKED (GLCNAC. . .).
CARBOHYD 94 94 N-LINKED (GLCNAC. . .).
CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
VARSPIC 236 237 GI -> VL (in isoform Short).
FTID-VSP_002522.
Missing (in isoform Short).
FTID-VSP_002523.
SEQUENCE 250 AA; 28147 MW; 34D635DFD1D14FE2E CRC64;
Query Match 100.0%; Score 327; DB 1; Length 250;
Best Local Similarity 100.0%; Pred. No. 5,9e-32;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LHCFCGIFSCSQIYGVYGVNTHVPSNVLKVKQKDKVAELENSEFRAFSFKN 60
|||||
20 LHCFCGIFSCSQIYGVYGVNTHVPSNVLKVKQKDKVAELENSEFRAFSFKN 79
|||||
61 R 61
80 R 80
QY Db QY Db
```

```
RESULT 2
ASPN_HUMAN
ID ASPN_HUMAN STANDARD; PRT; 379 AA.
AC Q9BXN1; Q96K79; Q96LDO;
DC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Asporin precursor (Periodontal ligament associated protein-1) (PLAP-1).
GN ASPN OR PLAP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., PARTIAL SEQUENCE, CARBOHYDRATE-LINKAGE SITE
RP ASN-281, MASS SPECTROMETRY, AND POLYMORPHISM OF POLY-ASP REGION.
RC TISSUE=Cartilage;
RX MEDLINE=21192276; PubMed=11152692;
RA Lorenzo P., Asperg A., Oennerfjord P., Bayliss M.T., Neame P.J.,
RA Heinegaard D.;
RT "Identification and characterization of asporin. A novel member of the
RT leucine-rich repeat protein family closely related to decorin and
RT biglycan."
RL J. Biol. Chem. 276:12201-12211(2001).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=21472263; PubMed=11587855;
RA Yamada S., Murakami S., Matoba K., Ozawa Y., Yokokoji T., Nakahira Y.,
RA Ikezawa K., Takayama S.-I., Matsubara K., Okada H.;
RT "Expression profile of active genes in human periodontal ligament and
RT isolation of PLAP-1, a novel SLRP family gene."
RL Gene 275:279-286(2001).
RN [3]
SEQUENCE OF 1-242 FROM N.A.
RC TISSUE=Embryo;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamanoto J., Wakamatsu A., Nakamura Y., Naganari K., Masuno Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NED0 human cDNA sequencing project."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=21192277; PubMed=11152695;
RA Henry S.P., Takanosu M., Boyd T.C., Mayne P.M., Eberspaecher H.,
RA Zhou W., de Crombrughe B., Hoebeek M., Mayne R.;
RT "Expression pattern and gene characterization of asporin. A newly
RT discovered member of the leucine-rich repeat protein family."
RL J. Biol. Chem. 276:12212-12221(2001).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- TISSUE SPECIFICITY: Higher levels in osteoarthritic articular
CC cartilage, aorta, uterus. Moderate expression in small intestine,
CC heart, liver, bladder, ovary, stomach, and in the adrenal,
CC thyroid, and mammary glands. Low expression in trachea, bone
CC marrow, and lung.
CC -!- PTM: Does not contain a serine/glycine dipeptide sequence required
CC for the assembly of O-linked glycosaminoglycans and is probably
CC not a proteoglycan. The N-linked glycan at Asp-281 is composed of
CC variable structures of GlcNAc, mannose, fucose, HexNAc and hexose.
CC -!- POLYMORPHISM: The poly-Asp region of ASPN is polymorphic and
CC ranges at least from 11 to 17 Asp.
CC -!- SIMILARITY: BELONGS TO THE SMALL LEUCINE-RICH PROTEOGLYCAN
CC (SLRP) FAMILY. CLASS I SUBFAMILY.
CC -!- SIMILARITY: Contains 12 leucine-rich (LRR) repeats.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
```

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M protein - protein search, using sw model

un on: October 7, 2003, 16:45:18 ; Search time 18.0907 seconds  
(without alignments)  
870.128 million cell updates/sec

title: US-09-730-465-2\_COPY\_20\_80

effect score: 327  
sequence: 1 LHCFGFISCSQIYGVVYG.....DKVAELENSEFRAPSSFKNR 61

coring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	327	100.0	240	4 Q9BRW0	Q9brw0 homo sapien
2	271	82.9	134	4 Q14748	Q14748 homo sapien
3	230	70.3	119	4 Q16393	Q16393 homo sapien
4	165.5	50.6	159	6 Q28752	Q28752 oviss sp. lf
5	165.5	50.6	227	6 Q28754	Q28754 oviss sp. lf
6	165.5	50.6	253	6 Q28753	Q28753 oviss sp. lf
7	111.5	34.1	244	6 Q8SQB6	Q8sqb6 sus scrofa
8	99	30.3	357	13 Q90ZL5	Q90zl5 anas platyr
9	67	20.5	546	5 Q9U2P8	Q9u2p8 caenorhabdi
10	65	19.9	108	2 Q8GN34	Q8gn34 uncultured
11	63.5	19.4	562	16 Q9KVG2	Q9kv22 vibrio chol
12	63	19.3	515	10 Q8SVG2	Q8sv26 oryza sativ
13	62.5	19.1	413	5 Q8I333	Q8i333 plasmodium
14	62.5	19.1	749	5 Q8IEC9	Q8iec9 plasmodium
15	62	19.0	205	5 Q15795	Q15795 plasmodium
16	62	19.0	252	5 Q8IC03	Q8ic03 plasmodium

17	62	19.0	416	5	Q18094	Q18094 caenorhabdi
18	62	19.0	1325	2	Q8GAT8	Q8gat8 alteromonas
19	61	18.7	416	5	Q9N507	Q9n507 caenorhabdi
20	61	18.7	1394	2	Q9AIU3	Q9aiu3 anaplasma p
21	60.5	18.5	275	16	Q9KGL6	Q9kg16 bacillus ha
22	60.5	18.5	662	5	P90663	P90663 aedes aegyp
23	60.5	18.5	903	16	Q92MI9	Q92mi9 rhizobium m
24	60.5	18.5	2924	5	Q25733	Q25733 plasmodium
25	60	18.3	619	10	Q9CAA5	Q9caa5 arabidopsis
26	60	18.3	852	16	Q8ZIT3	Q8zit3 versinia pe
27	60	18.3	878	16	Q8CZM8	Q8czm8 versinia pe
28	60	18.3	1342	5	Q8MMU8	Q8mmu8 dictyosteli
29	60	18.3	6751	5	Q8IC77	Q8ic77 plasmodium
30	59.5	18.2	273	16	Q25410	Q25410 helicobacte
31	59.5	18.2	288	6	Q77684	Q77684 macaca neme
32	59.5	18.2	288	6	Q28499	Q28499 macaca mula
33	59.5	18.2	458	16	Q8D2I6	Q8d2i6 wigglewort
34	59.5	18.2	734	5	Q8I1Z1	Q8i1z1 plasmodium
35	59.5	18.2	928	5	Q8IJN2	Q8ijn2 plasmodium
36	59.5	18.2	1141	6	Q46486	Q46486 pongo pygma
37	59	18.0	294	16	Q92B18	Q92b18 listeria in
38	59	18.0	351	5	Q9U0L5	Q9u0l5 plasmodium
39	59	18.0	387	10	Q8L8V2	Q8l8v2 arabidopsis
40	59	18.0	388	10	Q48727	Q48727 arabidopsis
41	59	18.0	404	10	Q8H091	Q8h091 oryza sativ
42	59	18.0	475	10	Q9STG5	Q9stg5 arabidopsis
43	59	18.0	495	10	Q8S7S4	Q8s7s4 oryza sativ
44	59	18.0	833	16	Q8ZKM7	Q8zkm7 salmonella
45	58.5	17.9	351	2	Q30720	Q30720 helicobacte

## ALIGNMENTS

### RESULT 1

Q9BRW0	PRELIMINARY;	PRT;	240 AA.
ID	Q9BRW0		
AC	Q9BRW0;		
DC	01-JUN-2001 (T-EMBLrel. 17, Created)		
DT	01-JUN-2001 (T-EMBLrel. 17, Last sequence update)		
DI	01-DEC-2001 (T-EMBLrel. 19, Last annotation update)		
DE	Similar to CD58 antigen, (Lymphocyte function-associated antigen 3).		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RA	Strausberg R.;		
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC005930; AAH05930.1; -		
DR	HSSP; P19256; ICCZ.		
DR	InterPro; IPR003599; Ig.		
DR	SMART; SM00409; IG; 1.		
SQ	SEQUENCE 240 AA; 27044 MW; 431E44EFDDF80B2 CRC64;		

Query Match 100.0%; Score 327; DB 4; Length 240;

Best Local Similarity 100.0%; Pred. No. 5,6e-32;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LHCFGFISCSQIYGVVYGNTFFHVPSPVLPKVKKKDKVAELENSEFRAPSSFKN 60

Db 20 LHCFGFISCSQIYGVVYGNTFFHVPSPVLPKVKKKDKVAELENSEFRAPSSFKN 79

Qy 61 R 61

Db 80 R 80

### RESULT 2

Q14748

[illegible]

RT "Anas platyrhynchos T cell antigens.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY032731; AAK51607.1; -.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR003006; Ig\_MHC.  
DR Pfam: PF00047; Ig; 2.  
DR SMART: SM00409; Ig; 1.  
SQ SEQUENCE 357 AA; 39930 MW; 274161DF83868C08 CRC64;

Query Match 30.3%; Score 99; DB 13; Length 357;  
Best Local Similarity 41.1%; Pred. No. 0.00074;  
Matches 23; Conservative 13; Mismatches 18; Indels 2; Gaps 2;

QY 7 ISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAELE-NSEFRAPSSPKNR 61  
DB 22 IYC-BEKVFGILGENTFFKADNKIDEVITKDKVAEWAQSEPTFYFNLSR 76

RESULT 9  
Q902P8  
ID Q902P8 PRELIMINARY; PRT; 546 AA.  
AC Q902P8;  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Y26D4A.10 protein.  
GN Y26D4A.10  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA White S.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C. elegans: A platform for  
investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL: AL110478; CAB54346.1; -.  
DR WormPep: Y26D4A.10; CE21490.  
SQ SEQUENCE 546 AA; 63551 MW; 191A4A7B76B06261 CRC64;

Query Match 20.5%; Score 67; DB 5; Length 546;  
Best Local Similarity 28.4%; Pred. No. 9.9;  
Matches 21; Conservative 11; Mismatches 26; Indels 16; Gaps 3;

QY 2 HCGFTSCPSQIIGVYVY-----GNVTFHVPSPVLPKLVKKQ--KDKV---A 45  
DB 165 HIFGFVECYNKIIYGRKHNLSYRSPENKNSITSSNPQNTAKVINKLOWIPIPL 224  
QY 46 ELENSEFRAPSSPK 59  
DB 225 EMONKLINIFESNK 238

RESULT 10  
Q8GN34  
ID Q8GN34 PRELIMINARY; PRT; 108 AA.  
AC Q8GN34;  
DT 01-MAR-2003 (TREMELrel. 23, Created)  
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE Tetraacycline resistance protein Tet37.  
GN Tet37  
OS uncultured bacterium.  
OC Bacteria; environmental samples.  
OX NCBI\_TaxID=77133;  
RN [1]  
RP SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.  
RA Kakutani T.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: D28584; BA005920.1; -.  
DR HSP: P19256; LCCZ.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR001993; Mitoch\_carrier.  
DR SMART: SM00409; Ig; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
FT SIGNAL.  
KW SIGNAL.  
SQ SEQUENCE 253 AA; 27715 MW; 4A4E97A38519B290 CRC64;

Query Match 50.6%; Score 165.5; DB 6; Length 253;  
Best Local Similarity 57.9%; Pred. No. 3.6e-12;  
Matches 33; Conservative 6; Mismatches 17; Indels 1; Gaps 1;

QY 6 FISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAEL-ENSEFRAPSSPKNR 61  
DB 25 FISCVSQDIAGMNGNVTFIVSESPFTEIMWKKGVNDVOTSGLEAFQSPKNR 81

RESULT 7  
Q8SQB6  
ID Q8SQB6 PRELIMINARY; PRT; 244 AA.  
AC Q8SQB6;  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE CD58 antigen.  
GN CD58.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Brossay A., Hube F., Bardos P., Watier H.;  
RT "Characterization of the porcine CD58 antigen mRNA sequence.";  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF469666; AAL76278.1; -.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR001993; Mitoch\_carrier.  
DR SMART: SM00409; Ig; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
SQ SEQUENCE 244 AA; 27301 MW; 164B81ACC7A512D2 CRC64;

Query Match 34.1%; Score 111.5; DB 6; Length 244;  
Best Local Similarity 42.1%; Pred. No. 1.4e-05;  
Matches 24; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 6 FISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAELENS-EFRAPSSPKNR 61  
DB 25 FICDCSVIYGLHKNVTLTRASSQTFQLIWKKGKDKAVWEQYNTATPPFVDR 81

RESULT 8  
Q90ZL5  
ID Q90ZL5 PRELIMINARY; PRT; 357 AA.  
AC Q90ZL5;  
DT 01-DEC-2001 (TREMELrel. 19, Created)  
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE CD58 antigen.  
OS Anas platyrhynchos (Domestic duck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
OX NCBI\_TaxID=8839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: D28584; BA005920.1; -.  
DR HSP: P19256; LCCZ.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR001993; Mitoch\_carrier.  
DR SMART: SM00409; Ig; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
FT SIGNAL.  
KW SIGNAL.  
SQ SEQUENCE 253 AA; 27715 MW; 4A4E97A38519B290 CRC64;

Query Match 50.6%; Score 165.5; DB 6; Length 253;  
Best Local Similarity 57.9%; Pred. No. 3.6e-12;  
Matches 33; Conservative 6; Mismatches 17; Indels 1; Gaps 1;

QY 6 FISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAEL-ENSEFRAPSSPKNR 61  
DB 25 FISCVSQDIAGMNGNVTFIVSESPFTEIMWKKGVNDVOTSGLEAFQSPKNR 81

RESULT 7  
Q8SQB6  
ID Q8SQB6 PRELIMINARY; PRT; 244 AA.  
AC Q8SQB6;  
DT 01-JUN-2002 (TREMELrel. 21, Created)  
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE CD58 antigen.  
GN CD58.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Brossay A., Hube F., Bardos P., Watier H.;  
RT "Characterization of the porcine CD58 antigen mRNA sequence.";  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF469666; AAL76278.1; -.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR001993; Mitoch\_carrier.  
DR SMART: SM00409; Ig; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
SQ SEQUENCE 244 AA; 27301 MW; 164B81ACC7A512D2 CRC64;

Query Match 34.1%; Score 111.5; DB 6; Length 244;  
Best Local Similarity 42.1%; Pred. No. 1.4e-05;  
Matches 24; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 6 FISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAELENS-EFRAPSSPKNR 61  
DB 25 FICDCSVIYGLHKNVTLTRASSQTFQLIWKKGKDKAVWEQYNTATPPFVDR 81

RESULT 8  
Q90ZL5  
ID Q90ZL5 PRELIMINARY; PRT; 357 AA.  
AC Q90ZL5;  
DT 01-DEC-2001 (TREMELrel. 19, Created)  
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE CD58 antigen.  
OS Anas platyrhynchos (Domestic duck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anas.  
OX NCBI\_TaxID=8839;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chan S.W.S., Middleton D.L., Lundqvist M., Warr G.W., Higgins D.A.;  
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: D28584; BA005920.1; -.  
DR HSP: P19256; LCCZ.  
DR InterPro: IPR003599; Ig.  
DR InterPro: IPR001993; Mitoch\_carrier.  
DR SMART: SM00409; Ig; 1.  
DR PROSITE: PS00215; MITOCH\_CARRIER; 1.  
FT SIGNAL.  
KW SIGNAL.  
SQ SEQUENCE 253 AA; 27715 MW; 4A4E97A38519B290 CRC64;

Query Match 50.6%; Score 165.5; DB 6; Length 253;  
Best Local Similarity 57.9%; Pred. No. 3.6e-12;  
Matches 33; Conservative 6; Mismatches 17; Indels 1; Gaps 1;

QY 6 FISCFSQIIGVYGVVGNVTFHVPSPVLPKLVKKQKDKVAEL-ENSEFRAPSSPKNR 61  
DB 25 FISCVSQDIAGMNGNVTFIVSESPFTEIMWKKGVNDVOTSGLEAFQSPKNR 81

RESULT 7  
Q8SQB6  
ID Q8SQB6 PRELIMINARY; PRT; 244 AA.  
AC Q8SQB6;  
DT 0

[illegible]

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:47:23 ; Search time 6.1299 Seconds  
(without alignments)  
421.045 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_20\_80

Perfect score: 327

Sequence: 1 LHCFGFISCFSQIYGVYG.....DKVAELENSEFRAFSSFKNR 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
  - 2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
  - 3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*
  - 4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/1aa/PCURS\_COMB.pep.\*
  - 6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	240	1	US-07-940-861-12
2	327	100.0	240	1	US-08-459-512-12
3	327	100.0	240	2	US-08-459-512-12
4	327	100.0	240	2	US-08-460-132-12
5	327	100.0	240	3	US-08-466-465-4
6	327	100.0	240	5	PCT-US92-02050-12
7	327	100.0	240	6	5185441-36
8	327	100.0	240	6	5223394-4
9	327	100.0	240	6	5223394-6
10	327	100.0	250	1	US-07-940-861-10
11	327	100.0	250	1	US-08-459-512-10
12	327	100.0	250	2	US-08-459-657-10
13	327	100.0	250	2	US-08-460-132-10
14	327	100.0	250	3	US-08-466-465-2
15	327	100.0	250	5	PCT-US92-02050-10
16	327	100.0	250	6	5223394-1
17	327	100.0	251	6	5185441-38
18	327	100.0	347	1	US-07-940-861-43
19	327	100.0	347	1	US-08-459-512-43
20	327	100.0	347	2	US-08-459-657-43
21	327	100.0	347	2	US-08-460-132-43
22	327	100.0	347	3	US-08-466-465-8
23	327	100.0	347	5	PCT-US92-02050-43
24	316	96.6	77	1	US-07-940-861-5
25	316	96.6	77	1	US-08-459-512-5
26	316	96.6	77	2	US-08-459-657-5
27	316	96.6	77	2	US-08-460-132-5

Sequence 5, Appli  
Sequence 13, Appl  
Sequence 8, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Patent No. 5185441  
Sequence 1, Appl  
Sequence 36, Appl  
Sequence 31, Appl  
Sequence 3, Appl  
Sequence 43, Appl  
Sequence 14, Appl  
Sequence 156, Appl  
Sequence 24, Appl  
Sequence 14, Appl

28 316 96.6 77 5 PCT-US92-02050-5  
29 271 82.9 134 1 US-08-328-152A-13  
30 271 82.9 222 1 US-08-328-152A-6  
31 260 79.5 50 1 US-07-940-861-2  
32 260 79.5 50 1 US-08-459-512-2  
33 260 79.5 50 2 US-08-459-657-2  
34 260 79.5 50 2 US-08-460-132-2  
35 260 79.5 50 5 PCT-US92-02050-2  
36 200 61.2 38 6 5185441-1  
37 143.5 43.9 131 1 US-08-328-152A-1  
38 143.5 43.9 199 1 US-08-328-152A-36  
39 143.5 43.9 225 1 US-08-328-152A-31  
40 62 19.0 29 1 US-08-328-152A-3  
41 60.5 18.5 373 4 US-09-724-864-43  
42 59.5 18.2 288 4 US-09-651-200-14  
43 57.5 17.6 272 4 US-09-325-932A-156  
44 56.5 17.3 179 4 US-09-513-057C-24  
45 56.5 17.3 317 4 US-08-913-159-14

#### ALIGNMENTS

RESULT 1  
US-07-940-861-12  
; Sequence 12, Application US/07940861  
; Patent No. 5547853  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/940,861  
; FILING DATE: 21-OCT-1992  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B1501P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear



MOLECULE TYPE: protein

US-07-940-861-12

Query Match 100.0%; Score 327; DB 1; Length 240;

Best Local Similarity 100.0%; Pred. No. 1.1e-34;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 60

20 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 79

61 R 61

80 R 80

RESULT 2

US-08-459-512-12

; Sequence 12, Application US/08459512

; Patent No. 5728677

; GENERAL INFORMATION:

; APPLICANT: BIOGEN, INC.

; APPLICANT: WALLNER, Barbara P.

; APPLICANT: MILLER, Glenn T.

; APPLICANT: ROSA, Margaret D.

; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Neave

; STREET: 875 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10022-6250

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,512

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02050

; FILING DATE: 12-MAR-1992

; APPLICATION NUMBER: US 07/667,971

; FILING DATE: 12-MAR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/770,967

; FILING DATE: 07-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, James F., Jr.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: B151CIP2

; TELEPHONE: (212)715-0600

; TELEFAX: (212)715-0673

; TELEX: 14-8367

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 240 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-459-512-12

Query Match

Best Local Similarity 100.0%; Score 327; DB 1; Length 240;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 60

Db 20 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 79

QY 61 R 61

80 R 80

RESULT 3

US-08-459-657-12

; Sequence 12, Application US/08459657

; Patent No. 5914111

; GENERAL INFORMATION:

; APPLICANT: BIOGEN, INC.

; APPLICANT: WALLNER, Barbara P.

; APPLICANT: MILLER, Glenn T.

; APPLICANT: ROSA, Margaret D.

; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Neave

; STREET: 875 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10022-6250

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,657

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02050

; FILING DATE: 12-MAR-1992

; APPLICATION NUMBER: US 07/667,971

; FILING DATE: 12-MAR-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/770,967

; FILING DATE: 07-OCT-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: HALEY, James F., Jr.

; REGISTRATION NUMBER: 27,794

; REFERENCE/DOCKET NUMBER: B151CIP2

; TELEPHONE: (212)715-0600

; TELEFAX: (212)715-0673

; TELEX: 14-8367

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 240 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-459-657-12

Query Match

Best Local Similarity 100.0%; Score 327; DB 2; Length 240;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 60

Db 20 LHCFGFISCFSSQIYGVVGNVTFHVPSPVPLKEVLWKKQDKVAELENSEFRASFKN 79

QY 61 R 61

80 R 80

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:52:18 ; Search time 12.7083 Seconds  
(without alignments)  
759.422 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_20\_80

Perfect score: 327

Sequence: 1 LHCFGISCFSSQIIGVYGVG.....DKVAELENSEFRAPSFKNR 61

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries.

Database : Published Applications\_AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	240	9	US-09-796-033-4
2	327	100.0	240	9	US-09-730-465-4
3	327	100.0	250	9	US-09-796-033-2
4	327	100.0	250	9	US-09-730-465-2
5	327	100.0	250	12	US-10-436-523-61
6	327	100.0	250	15	US-10-207-655-113
7	327	100.0	347	9	US-09-796-033-8
8	327	100.0	347	9	US-09-730-465-8
9	327	100.0	347	14	US-10-091-236-17
10	327	100.0	347	15	US-10-091-313-7
11	327	100.0	347	15	US-10-091-268-7
12	62.5	19.1	287	12	US-10-032-214-237
13	61.5	18.8	379	9	US-09-866-028-2
14	61.5	18.8	379	10	US-09-944-449-2
15	61.5	18.8	379	10	US-09-944-457-2

16	61.5	18.8	379	10	US-09-944-862-2	Sequence 2, Appli
17	61.5	18.8	379	10	US-09-945-587-2	Sequence 2, Appli
18	61.5	18.8	379	10	US-09-945-015-2	Sequence 2, Appli
19	61.5	18.8	379	10	US-09-944-396-2	Sequence 2, Appli
20	61.5	18.8	379	10	US-09-944-097-2	Sequence 2, Appli
21	61.5	18.8	379	10	US-09-944-432-2	Sequence 2, Appli
22	61.5	18.8	379	10	US-09-943-762-2	Sequence 2, Appli
23	61.5	18.8	379	10	US-09-944-654-2	Sequence 2, Appli
24	61.5	18.8	379	10	US-09-943-851A-2	Sequence 2, Appli
25	61.5	18.8	379	10	US-09-944-413-2	Sequence 2, Appli
26	61.5	18.8	379	10	US-09-944-403-2	Sequence 2, Appli
27	61.5	18.8	379	10	US-09-944-896-2	Sequence 2, Appli
28	61.5	18.8	379	10	US-09-944-944-2	Sequence 2, Appli
29	61.5	18.8	379	10	US-09-944-929-2	Sequence 2, Appli
30	61.5	18.8	379	10	US-09-944-907-2	Sequence 2, Appli
31	61.5	18.8	379	11	US-09-944-884-2	Sequence 2, Appli
32	61.5	18.8	379	11	US-09-944-852-2	Sequence 2, Appli
33	61.5	18.8	379	11	US-09-943-780-2	Sequence 2, Appli
34	61.5	18.8	379	12	US-10-137-870-328	Sequence 328, App
35	61.5	18.8	379	12	US-10-140-018-328	Sequence 328, App
36	61.5	18.8	379	12	US-10-140-021-328	Sequence 328, App
37	61.5	18.8	379	12	US-10-140-274-328	Sequence 328, App
38	61.5	18.8	379	12	US-10-140-471-328	Sequence 328, App
39	61.5	18.8	379	12	US-10-140-807-328	Sequence 328, App
40	61.5	18.8	379	12	US-10-140-922-328	Sequence 328, App
41	61.5	18.8	379	12	US-10-140-924-328	Sequence 328, App
42	61.5	18.8	379	12	US-10-140-926-328	Sequence 328, App
43	61.5	18.8	379	12	US-10-141-698-328	Sequence 328, App
44	61.5	18.8	379	12	US-10-141-702-328	Sequence 328, App
45	61.5	18.8	379	12	US-10-141-704-328	Sequence 328, App

#### ALIGNMENTS

RESULT 1  
US-09-796-033-4  
; Sequence 4, Application US/09796033  
; Patent No. US20020009446A1  
; GENERAL INFORMATION:  
; APPLICANT: Megilavy, Daniel  
; TITLE OF INVENTION: METHOD OF MODULATING MEMORY EFFECTOR  
; FILE REFERENCE: T-CELLS AND COMPOSITIONS  
; FILE REFERENCE: 10274-044001  
; CURRENT APPLICATION NUMBER: US/09/796,033  
; CURRENT FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: PCT/US99/20026  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: US 60/098,456  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIGNAL  
; LOCATION: (1)...(28)  
US-09-796-033-4

Query Match 100.0%; Score 327; DB 9; Length 240;  
Best Local Similarity 100.0%; Pred. No. 3.8e-34;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LHCFGISCFSSQIIGVYGVYGVNTPFVPSNPVLKYLWKQKDKVAELENSEFRAPSFKN 60

Db 20 LHCFGISCFSSQIIGVYGVYGVNTPFVPSNPVLKYLWKQKDKVAELENSEFRAPSFKN 79

Qy 61 R 61

Db 80 R 80

```

; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: PCT/US99/20026
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: US 60/098,456
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(28)
US-09-796-033-2

Query Match      100.0%; Score 327; DB 9; Length 250;
Best Local Similarity 100.0%; Pred. No. 4e-34;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LHCFGTSCFSQIYGVYGVNTHFVPSNPLKEVLWKOKDKVAELENSEFFRAFSSFKN 60
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      20 LHCFGTSCFSQIYGVYGVNTHFVPSNPLKEVLWKOKDKVAELENSEFFRAFSSFKN 79
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      61 R 61
        |
Db      80 R 80

RESULT 4
US-09-730-465-2
; Sequence 2, Application US/09730465
; Patent No. US2002009449A1
; GENERAL INFORMATION:
; APPLICANT: Wallner, Barbara P.
; TITLE OF INVENTION: Method of Prophylaxis or Treatment of Antigen
; Presenting Cell Driven Skin Conditions Using
; Inhibitors of the CD2/LFA-3 Interaction
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/730,465
; FILING DATE: 05-Dec-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08755
; FILING DATE: 06-OCT-1992
; APPLICATION NUMBER: US 07/862,022
; FILING DATE: 12-APR-1992
; APPLICATION NUMBER: US 07/770,969
; FILING DATE: 07-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis (PLM)
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: BGF-111CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:24:22 ; Search time 81 Seconds  
(without alignments)  
31.353 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_50\_65  
Sequence: 1 PLKEVLWKKQDKVAE 16

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 360314

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

#### Database :

A\_Geneseq\_19Jun03:\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT:\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT:\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT:\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT:\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT:\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT:\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT:\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT:\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT:\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT:\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT:\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT:\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT:\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT:\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT:\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT:\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT:\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT:\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT:\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT:\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT:\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT:\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT:\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	42.9	16	23	AAE23846 Human zsig33-gamma
2	36	42.9	16	23	AAE23847 Human zsig33-gamma
3	36	42.9	16	23	AAE15891 Human zsig33-gamma
4	36	42.9	16	23	AAE15892 Human zsig33-gamma
5	35	41.7	15	16	AAE73948 Fragment of T.sery
6	35	41.7	15	23	ABF58879 Tissue anion trans
7	34	40.5	10	21	AAE69683 Human c-Erb-A nucl
8	34	40.5	15	23	AAE23848 Human zsig33-gamma
9	34	40.5	15	23	AAE15893 Human zsig33-gamma

10	34	40.5	15	24	ABR33624 Human cancer-relat
11	34	40.5	15	24	ABR33657 Human cancer-relat
12	34	40.5	15	24	ABR33676 Human cancer-relat
13	34	40.5	15	24	ABR33694 Human cancer-relat
14	34	40.5	15	24	ABR33717 Human cancer-relat
15	34	40.5	15	24	ABR33753 Human cancer-relat
16	34	40.5	15	24	ABR33764 Human cancer-relat
17	32	38.1	14	22	AAE98523 Human peptide #179
18	32	38.1	16	15	AAE66027 Pair 3 pep 5 immun
19	31	36.9	8	23	AAU76075 Nociceptin-like im
20	31	36.9	9	23	AAU76074 Nociceptin-like im
21	31	36.9	10	22	AAE31762 Amino acid sequenc
22	31	36.9	10	23	ABP53304 Gene delivery rela
23	31	36.9	11	23	AAU76110 Nociceptin-like pe
24	31	36.9	12	23	AAU76099 Nociceptin-like pe
25	31	36.9	12	23	AAU76102 Nociceptin-like pe
26	31	36.9	12	23	AAU76106 Nociceptin-like pe
27	31	36.9	12	23	AAU76107 Nociceptin-like pe
28	31	36.9	12	23	AAU76112 Nociceptin-like pe
29	31	36.9	13	23	AAU76111 Nociceptin-like pe
30	31	36.9	14	18	AAW24446 Nucleic acid (NA)
31	31	36.9	14	22	AAE96920 Human peptide #195
32	31	36.9	14	22	AAE00706 Human protein frag
33	31	36.9	15	23	ABP53305 Gene delivery rela
34	31	36.9	16	21	AAV52664 yes-encoded oncopr
35	30	35.7	8	18	AAW44604 Anti-fungal peptid
36	30	35.7	8	20	AAE00581 Anti-fungal peptid
37	30	35.7	8	22	AAE65505 Anti-fungal peptid
38	30	35.7	9	17	AAW04048 Anti-fungal peptid
39	30	35.7	9	18	AAW44562 Anti-fungal peptid
40	30	35.7	9	20	AAE00539 Anti-fungal peptid
41	30	35.7	9	22	AAE65463 Anti-fungal peptid
42	30	35.7	14	14	AAE31887 Protein S - C4bp b
43	30	35.7	16	15	AAE47955 CytLA-2 beta fragme
44	30	35.7	16	20	AAE42318 Synthetic cytotoxi
45	29	34.5	13	22	AAE03136 Human gene 7 encod

#### ALIGNMENTS

RESULT 1  
AAE23846  
ID AAE23846 standard; peptide; 16 AA.  
AC AC  
XX AAE23846;  
XX  
DT 10-SFP-2002 (first entry)  
XX  
DE Human zsig33-gamma peptide #2.  
XX  
KW Human; zsig33-like peptide; gastric contractility; nutrient uptake;  
KW growth hormone; digestive enzyme; restorative therapy; gene therapy;  
KW protein therapy; gastrointestinal; endocrine; anabolic.  
XX Homo sapiens.  
XX  
PN US200205156-A1.  
XX  
PD 09-MAY-2002.  
XX  
PF 10-MAY-2001; 2001US-0853253.  
XX  
PR 11-MAY-2000; 2000US-203300P.  
XX  
PA (JASP/) JASPERS S R.  
PA (SHEP/) SHEPPARD P O.  
PA (DEIS/) DEISHER T A.  
PA (BISH/) BISHOP P D.  
XX Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;  
XX WPI; 2002-443750/47.

ZSIG33-Like peptides and polynucleotides, useful for modulating gastric contractility, nutrient uptake, growth hormones and/or secretion of digestive/pancreatic enzymes and hormones -

Claim 5; Page 30; 34pp; English.

The invention relates to zsig33-like peptides and their corresponding nucleic acids and methods for modulating gastric contractility, nutrient uptake, growth hormones, secretion of digestive enzymes and hormones. The sequences of the invention are used in the prevention, diagnosis and treatment of diseases associated with inappropriate ZSIG33 expression. The nucleic acids of the invention and their complements are used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The ZSIG33 peptides are used as antigens in the production of antibodies against zsig33 and in assays to identify modulators of ZSIG33 expression and activity. The anti-ZSIG33 antibodies and antagonists are used to down regulate expression and activity. The anti-ZSIG33 antibodies are also used as diagnostic agents for detecting the presence of ZSIG33 in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The peptides and nucleic acids of the invention are used to modulate gastric contractility, nutrient uptake, growth hormones, the secretion of digestive enzymes and hormones, and/or secretion of enzymes and/or hormones in the pancreas. zsig33-like DNA is used in gene therapy and zsig33-like peptide is used in protein therapy. The present sequence is human zsig33-like peptide, zsig33-gamma peptide.

Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;  
Best Local Similarity 33.3%; Pred. No. 91;  
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQKDKVAE 16  
|:::||::||:|:  
DB 2 LQDILWEEAKEAPAD 16

RESULT 2

AAP23847

ID AAE223847 standard; peptide; 16 AA.  
XX RAE23847;  
AC AC  
CC CC  
DT DT  
XX XX  
DE DE  
XX XX  
KW Human; zsig33-like peptide; gastric contractility; nutrient uptake; growth hormone; digestive enzyme; restorative therapy; gene therapy; protein therapy; gastrointestinal; endocrine; anabolic.  
KW KW  
XX XX  
OS Homo sapiens.  
XX XX  
PN US2002055156-A1.  
XX XX  
PD PD  
XX 09-MAY-2002.  
XX XX  
PF 10-MAY-2001; 2001US-0853253.  
XX XX  
PR 11-MAY-2000; 2000US-203300P.  
XX XX  
PA (JASP/) JASPERS S R.  
XX PA (SHEP/) SHEPPARD P O.  
PA (DEIS/) DEISHER T A.  
PA (BISH/) BISHOP P D.  
XX XX  
PI Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;  
XX WPT; 2002-443750/47.  
XX XX  
PT ZSIG33-Like peptides and polynucleotides, useful for modulating gastric

contractility, nutrient uptake, growth hormones and/or secretion of digestive/pancreatic enzymes and hormones -

Claim 5; Page 30; 34pp: English.

The invention relates to zsig33-like peptides and their corresponding nucleic acids and methods for modulating gastric contractility, nutrient uptake, growth hormones, secretion of digestive enzymes and hormones. The sequences of the invention are used in the prevention, diagnosis and treatment of diseases associated with inappropriate zsig33 expression. The nucleic acids of the invention and their complements are used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. The zsig33 peptides are used as antigens in the production of antibodies against zsig33 and in assays to identify modulators of zsig33 expression and activity. The anti-zsig33 antibodies and antagonists are used to down regulate expression and activity. The anti-zsig33 antibodies are also used as diagnostic agents for detecting the presence of zsig33 in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The peptides and nucleic acids of the invention are used to modulate gastric contractility, nutrient uptake, growth hormones, the secretion of digestive enzymes and hormones, and/or secretion of enzymes and/or hormones in the pancreas. zsig33-like DNA is used in gene therapy and zsig33-like peptide is used in protein therapy. The present sequence is human zsig33-like peptide, zsig33-gamma peptide.

Sequence 16 AA;

Query Match 42.9%; Score 36; DB 23; Length 16;  
Best Local Similarity 33.3%; Pred. No. 91;  
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Oy 2 LKEVLWKQKDKVAE 16  
1:::ll::l: l:  
Db 2 LQDILWEAKEAPAD 16

RESULT 3

AAE15891

ID AAE15891 standard; peptide; 16 AA.

XX AAE15891;

XX 26-MAR-2002 (first entry)

XX Human zsig33-gamma peptide #2.

DE Human: zsig33-like peptide; zsig33LP: immunity; developmental process;  
KW infection; human immunodeficiency virus; vaccine; antihypoglycaemic;  
KW adsorption enhancer; gastrointestinal disease; growth related disease;  
KW inflammation; gene therapy; growth regulation; blood vessel formation;  
KW HIV; zsig33-gamma peptide.

XX Homo sapiens.

XX WO200187933-A2.

XX 22-NOV-2001.

XX 10-MAY-2001; 2001WO-US15091.

XX 11-MAY-2000; 2000US-0569271.

XX (ZTMO ) ZYMOGENETICS INC.

XX Jaspers SR, Sheppard PO, Deisher TA, Bishop PD;

XX WPI; 2002-082982/11.

DR N-PSDB; AAD25764.

XX New polypeptides, useful for modulating gastric contractility, nutrient uptake, pancreatic secretion of hormones, digestive enzymes and treating gastrointestinal and growth related diseases, comprises

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:25:08 ; Search time 22 Seconds  
(without alignments)  
34,201 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_50\_65  
Perfect score: 84  
Sequence: 1 PLKEVLWKKQDKVAE 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 872

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	23	27.4	15	1	MALT_BACTQ
2	21	25.0	13	1	NO40_PEA
3	21	25.0	13	1	NO40_VICSA
4	21	25.0	13	1	RPOC_MVCGA
5	21	25.0	14	1	MAST_PARID
6	21	25.0	14	1	MAST_POLJA
7	20	23.8	12	1	UR2B_CATCO
8	20	23.8	12	1	UR2B_CTPCA
9	20	23.8	12	1	UR2_POLSP
10	20	23.8	14	1	SMS1_MYOSC
11	20	23.8	14	1	SMS_ALIMI
12	18.5	22.0	14	1	MAST_VESBA
13	18	21.4	10	1	SPI_HALRO
14	18	21.4	11	1	LPW_THETH
15	18	21.4	12	1	NO40_SESRO
16	18	21.4	15	1	PGKH_PHIPA
17	18	21.4	16	1	FIBA_MELME
18	18	21.4	16	1	FIBA_MUSVI
19	17	20.2	7	1	WMA2_ACHFU
20	17	20.2	7	1	WMA3_ACHFU
21	17	20.2	9	1	LPCA_STAAR
22	17	20.2	11	1	CEP1_ACHFU
23	17	20.2	11	1	Q20A_COMTE
24	17	20.2	12	1	NO40_LOTJA
25	17	20.2	12	1	RS19_CLYEP
26	17	20.2	12	1	RS19_TOBBP
27	17	20.2	12	1	UR2A_CATCO
28	17	20.2	12	1	UR2_GILMI
29	17	20.2	12	1	UR2_SVCVA
30	17	20.2	13	1	AU11_LITRA
31	17	20.2	13	1	AU12_LITRA
32	17	20.2	14	1	RS19_PRAAP
33	17	20.2	15	1	C10A_RAT

34	17	20.2	15	1	CYSK_CLOPA	P81340 clostridium
35	17	20.2	15	1	THL_CLOPA	P81347 clostridium
36	17	20.2	15	1	UC27_MAIZE	P80633 zea mays (m
37	17	20.2	16	1	MMPX_SOLTU	P80501 solanum tub
38	16	19.0	9	1	COW_CONVE	P83047 conus ventr
39	16	19.0	9	1	NEF_HVIZ8	P12481 human immun
40	16	19.0	9	1	ULAD_HUMAN	P31929 homo sapien
41	16	19.0	10	1	APE_CAFGI	P80474 capocytoph
42	16	19.0	10	1	GONL_PETMA	P04378 petromyzon
43	16	19.0	10	1	GON2_PETMA	P30948 petromyzon
44	16	19.0	12	1	CALM_TETTH	Q08055 tetrahymena
45	16	19.0	12	1	NO40_SOYBN	P58960 glycine max

## ALIGNMENTS

RESULT 1  
MALT\_BACTQ STANDARD; PRT; 15 AA.  
AC P80072;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-DEC-1992 (Rel. 24, Last annotation update)  
DE Maltase (EC 3.2.1.20) (Alpha-glucosidase I) (Fragment).  
OS Bacillus thermoamyloliquefaciens.  
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1425;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=KP1071 / FERM P8477;  
RX MEDLINE=92209510; PubMed=1555585;  
RA Suzuki Y., Yonezawa K., Hattori M., Takii Y.;  
RT "Assignment of Bacillus thermoamyloliquefaciens KP1071  
alpha-glucosidase I to an exo-alpha-1,4-glucosidase, and its striking  
similarity to bacillary oligo-1,6-glucosidases in N-terminal sequence  
and in structural parameters calculated from the amino acid  
composition.";  
RL Eur. J. Biochem. 205:249-256(1992).  
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-  
linked D-glucose residues with release of D-glucose.  
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
DR PIR: S21240; S21240.  
KW Hydrolase; Glycosidase.  
FT NON\_TER 15 15  
SQ SEQUENCE 15 AA; 1929 MW; 62B4CE501F2D3042 CRC64;  
  
Query Match 27.4%; Score 23; DB 1; Length 15;  
Best Local Similarity 37.5%; Pred. No. 8.1e+02;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 2 LKEVLWKK 9  
Db :: ::  
1 MKRAWKE 8  
  
RESULT 2  
NO40\_PEA STANDARD; PRT; 13 AA.  
AC P55959;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Early nodulin 40.  
GN ENOD40.  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=cv. Sparkle; TISSUE=Root nodules;  
RX MEDLINE=95036021; PubMed=7948896;  
RA Matvienko M., van de Sande K., Yang W.-C., van Kammen A., Bisseling T.,  
FR Fransen H.J.;  
RT "Comparison of soybean and pea ENOD40 cDNA clones representing genes  
RT expressed during both early and late stages of nodule development.";  
RL Plant Mol. Biol. 26:487-493(1994).  
CC -!- FUNCTION: MODULATES THE ACTION OF AUXIN, AND MAY FUNCTION AS PLANT  
CC GROWTH REGULATOR THAT ALTERS PHYTOHORMONE RESPONSES (BY  
CC SIMILARITY).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE EARLY STAGES OF THE NODULE  
CC DEVELOPMENT.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X81064; -; NOT\_ANNOTATED\_CDS.  
DR Modulation.  
KW SEQUENCE 13 AA; 1565 MW; 3C695B6BD8A26C3 CRC64;  
SQ  
  
Query Match 25.0%; Score 21; DB 1; Length 13;  
Best Local Similarity 37.5%; Pred. No. 1.5e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 LKEVLMKK 9  
: : :  
: : :  
Db 1 MKFLCWQK 8  
  
RESULT 3  
NO40\_VICSA STANDARD; PRT; 13 AA.  
ID NO40\_VICSA  
AC P55961;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Early nodulin 40.  
EN ENOD40.  
OS Vicia sativa (Spring vetch) (Tare).  
GC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.  
OX NCBI\_TaxID=3908;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nigra; TISSUE=Root nodules;  
RX MEDLINE=96011756; PubMed=7548828;  
RA Vijn I., Yang W.-C., Pallisgaard N., Oestergaard Jensen E.,  
FR van Kammen A., Bisseling T.;  
RT "VsENOD5, VsENOD12 and VsENOD40 expression during Rhizobium-induced  
RT nodule formation on Vicia sativa roots.";  
RL Plant Mol. Biol. 28:1111-1119(1995).  
CC -!- FUNCTION: MODULATES THE ACTION OF AUXIN, AND MAY FUNCTION AS PLANT  
CC GROWTH REGULATOR THAT ALTERS PHYTOHORMONE RESPONSES (BY  
CC SIMILARITY).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED DURING RHIZOBIUM-INDUCED NODULE  
CC FORMATION. IN 4-DAY OLD NODULES IT IS FOUND IN ALL THE CELLS OF  
CC THE CENTRE OF THE NODULE PRIMORDIUM AND ALSO OCCURS IN THE REGION  
CC OF THE ROOT PERICYCLE FACING THE NODULE PRIMORDIUM. AT DAY 5,  
CC EXPRESSION IS SEEN IN THE COMPLETE CENTRAL TISSUE. AT DAY 20  
CC EXPRESSED IN THIS ZONE IT IS FOUND ONLY IN THE INFECTED CELLS BUT NOT  
CC IN THE UNINFECTED CELLS. AT THE TRANSITION OF PREFIXATION ZONE II  
CC INTO INTERZONE II-III EXPRESSION DECREASES IN THE INFECTED CELLS.  
CC IN THE FIXATION ZONE III, EXPRESSION IS INDUCED IN THE UNINFECTED  
CC CELLS AND IN THE PROXIMAL PART OF THIS ZONE IT IS UNDETECTABLE.  
CC PRESENT AT HIGH LEVELS IN THE PERICYCLE OF THE NODULE VASCULAR  
CC BUNDLE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X83683; CAB37926.1; -;  
DR PIR: S60046; S60046.  
KW Modulation.  
SQ SEQUENCE 13 AA; 1531 MW; 3C6953C4BD8A26C3 CRC64;  
  
Query Match 25.0%; Score 21; DB 1; Length 13;  
Best Local Similarity 37.5%; Pred. No. 1.5e+03;  
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 LKEVLMKK 9  
: : :  
: : :  
Db 1 MKFLCWQK 8  
  
RESULT 4  
RPOC\_MYCGA STANDARD; PRT; 13 AA.  
ID RPOC\_MYCGA  
AC P47716;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase  
DE beta' chain) (RNA polymerase beta' subunit) (Fragment).  
GN RPOC.  
OS Mycoplasma cellisepticum.  
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AS969Var.B;  
RA Skamrov A.V., Feoktistova E.S., Gol'dman M.A., Feoktistova E.S.,  
RA Bibilashvili R.S.;  
RL Submitted (XXX-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
CC BETA' CHAIN.  
CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L38402; AAB40952.1; -;  
DR Transferase; DNA-directed RNA polymerase; Transcription.  
FT NONTER 13 13  
SQ SEQUENCE 13 AA; 1630 MW; 4BEC27C7480D4333 CRC64;  
  
Query Match 25.0%; Score 21; DB 1; Length 13;  
Best Local Similarity 66.7%; Pred. No. 1.5e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 8 KKKQDK 13  
: : :  
: : :  
Db 6 KKKNK 11

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:25:43 ; Search time 92 Seconds  
(without alignments)  
44.879 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_50\_65  
Perfect score: 84  
Sequence: 1 PLKEVLWKQKDKVAE 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 4022

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	34.5	15	5 P82207	P82207 bombyx mori
2	27	32.1	14	2 Q47599	Q47599 escherichia
3	26	31.0	15	4 Q9UC7	Q9UC7 homo sapien
4	25	29.8	15	3 Q9URE0	Q9URE0 saccharomyc
5	24.5	29.2	13	2 Q9R8R9	Q9R8R9 streptococc
6	23	27.4	12	11 Q8CJ30	Q8CJ30 mus musculu
7	23	27.4	15	2 Q9R5L9	Q9R5L9 bacillus st
8	23	27.4	15	10 Q9SRV2	Q9SRV2 triticum ae
9	22	26.2	12	5 Q81574	Q81574 osteragia
10	21	25.0	12	10 Q93WF2	Q93WF2 lupinus lut
11	21	25.0	15	2 Q9R586	Q9R586 serratia ma
12	21	25.0	15	2 Q05991	Q05991 staphylococ
13	21	25.0	15	6 Q9GLJ0	Q9GLJ0 sus scrofa
14	21	25.0	16	4 Q81WX4	Q81WX4 homo sapien
15	21	25.0	16	11 Q9CTB2	Q9CTB2 mus musculu
16	20	23.8	9	4 Q9UNAO	Q9UNAO homo sapien

17	20	23.8	11	2 Q56413	Q56413 escherichia
18	20	23.8	11	4 Q9H4H5	Q9H4H5 homo sapien
19	20	23.8	11	4 Q9UE69	Q9UE69 homo sapien
20	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
21	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
22	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
23	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
24	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
25	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
26	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
27	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
28	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
29	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
30	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
31	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
32	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
33	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
34	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
35	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
36	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
37	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
38	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
39	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
40	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
41	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
42	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
43	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
44	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca
45	20	23.8	12	13 Q8JHCL	Q8JHCL cyprinus ca

#### ALIGNMENTS

RESULT 1  
ID P82207 PRELIMINARY; PRT; 15 AA.  
AC P82207;  
DT 01-OCT-2001 (TREMREL. 18, Created)  
DT 01-OCT-2001 (TREMREL. 18, Last sequence update)  
DE Unknown protein from 2D-page (Fragment).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]  
RP SEQUENCE.  
RC STRAIN=XINHANG X KEMING; TISSUE=Body wall, and Fat body;  
RX MEDLINE=21177481; PubMed=11280994;  
RA Zhong B.X.;  
RT "protein database for several tissues derived from five instar of  
silkworm";  
RL I Chuan Hsueh Pao 28:217-224(2001).  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 1877 MW; 580F6BD4703CA70C CRC64;  
Query Match 34.5%; Score 29; DB 5; Length 15;  
Best Local Similarity 35.7%; Pred. No. 6.6e+02;  
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
QY 3 KEVLWKQKDKVAE 16  
DB 2 KSLFTQKQVDNINE 15  
RESULT 2  
ID Q47599 PRELIMINARY; PRT; 14 AA.  
AC Q47599;  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)



01-NOV-1998 (TReMBLrel. 08, Last annotation update)  
C (Fragment).  
Escherichia coli.  
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
NCBI\_TaxID=562;  
[1]  
RN  
SEQUENCE FROM N.A.  
MEDLINE=91139577; PubMed=1995588;  
Tao T., Bourne J.C., Blumenthal R.M.;  
"A family of regulatory genes associated with type II restriction-  
modification systems.";  
J. Bacteriol. 173:1367-1375(1991).  
J. Bacteriol. 173:1367-1375(1991).  
EMBL: M63619; AAA24555.1; -;  
NON\_TER 1  
FT  
SEQUENCE 14 AA; 1705 MW; 77B6CA60581A4F3B CRC64;  
SQ  
Query Match 32.1%; Score 27; DB 2; Length 14;  
Best Local Similarity 45.5%; Pred. No. 1.3e+03;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
2Y 2 LKEVLWKQKD 12  
||||:|:|:  
Db 1 LKEVIMKKHE 11  
||||:|:|:  
RESULT 3  
Q9UC77 PRELIMINARY; PRT; 15 AA.  
ID Q9UC77  
AC Q9UC77  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Midline (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
RN  
SEQUENCE  
MEDLINE=94059921; PubMed=82411100;  
Novotny W.F., Maifi T., Mehta R.L., Milner P.G.;  
"Identification of novel heparin-releasable proteins, as well as the  
cytokines midline and pleiotrophin, in human postheparin plasma.";  
Arterioscler. Thromb. 13:1798-1805(1993).  
SEQUENCE 15 AA; 1527 MW; C34B6B97878474AC CRC64;  
SQ  
Query Match 31.0%; Score 26; DB 4; Length 15;  
Best Local Similarity 83.3%; Pred. No. 2e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 9 KOKDXV 14  
|:|||||  
Db 3 KKKDXV 8  
|:|||||  
RESULT 4  
Q9URE0 PRELIMINARY; PRT; 15 AA.  
ID Q9URE0  
AC Q9URE0  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE Type II topoisomerase, topoisomerase II (Fragment).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
NCBI\_TaxID=4932;  
[1]  
RN  
SEQUENCE FROM N.A.  
MEDLINE=95130576; PubMed=7829529;  
Elsea S.H., Hsiung Y., Nitiss J.L., Osheroff N.;

J. Biol. Chem. 270:1913-1920(1995).  
NON\_TER 1  
FT NON\_TER 15  
SQ SEQUENCE 15 AA; 2094 MW; 0A6A37F6B81B85F6 CRC64;  
Query Match 29.8%; Score 25; DB 3; Length 15;  
Best Local Similarity 40.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;  
QY 7 WKQKDKVAE 16  
:|:|:|:|:  
Db 2 YQKRDYMS 11  
:|:|:|:|:  
RESULT 5  
Q9R8R9 PRELIMINARY; PRT; 13 AA.  
ID Q9R8R9  
AC Q9R8R9  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE C5a peptidase (Fragment).  
GN SCPA.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AF1;  
RX MEDLINE=98298075; PubMed=9632622;  
RA Berge A., Rasmussen M., Bjorck L.;  
"Identification of an insertion sequence located in a region encoding  
virulence factors of Streptococcus pyogenes.";  
RL Infect. Immun. 66:3449-3453(1998).  
DR EMBL: AF064540; AAC38768.1; -;  
FT NON\_TER 13  
SQ SEQUENCE 13 AA; 1603 MW; 5EDADCD6ACE6723 CRC64;  
Query Match 29.2%; Score 24.5; DB 2; Length 13;  
Best Local Similarity 63.6%; Pred. No. 3e+03; 0; Indels 3; Gaps 1;  
Matches 7; Conservative 1; Mismatches 0;  
QY 8 KQKQ---DKVA 15  
|||||:|:|:  
Db 3 KQKQLPDKLA 13  
|||||:|:|:  
RESULT 6  
Q8CJ30 PRELIMINARY; PRT; 12 AA.  
ID Q8CJ30  
AC Q8CJ30  
DT 01-MAR-2003 (TReMBLrel. 23, Created)  
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Myoneurin (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
RN  
SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Kidney;  
RX MEDLINE=20334280; PubMed=10873615;  
RA Alliel P.M., Seddigi N., Goudou D., Cifuentes-Diaz C., Romero N.,  
RA Velasco E., Rieger F., Perin J.P.;  
RT "Myoneurin, a novel member of the BTB/POZ-zinc finger family highly  
expressed in human muscle.";  
RL Biochem. Biophys. Res. Commun. 273:385-391(2000).  
[2]  
RN  
SEQUENCE FROM N.A.  
RC STRAIN=BALB/c; TISSUE=Kidney;  
RA Bitoun M., Perin J.P., Seddigi N., Goudou D., Camuzat A., Mattei M.G.;

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	42	50.0	7	6	5185441-12	Patent No. 5185441
2	31	36.9	10	4	US-09-050-811-5	Sequence 5, Appli
3	31	36.9	15	4	US-09-050-811-6	Sequence 6, Appli
4	31	36.9	15	4	US-09-050-811-11	Sequence 11, Appl
5	30	35.7	8	2	US-08-621-259A-205	Sequence 205, App
6	30	35.7	9	2	US-08-621-259A-163	Sequence 163, App
7	30	35.7	9	5	PCR-US95-09262-163	Sequence 163, App
8	30	35.7	14	1	US-07-907-190-13	Sequence 13, Appl
9	30	35.7	16	2	US-08-485-937-11	Sequence 11, Appl
10	30	35.7	16	2	US-08-373-213-11	Sequence 11, Appl
11	30	35.7	16	5	PCR-US93-06552-11	Sequence 11, Appl
12	29	34.5	15	1	US-08-311-611A-116	Sequence 116, App
13	29	34.5	15	1	US-08-372-783-116	Sequence 116, App
14	29	34.5	15	1	US-08-372-105-116	Sequence 116, App
15	29	34.5	15	1	US-08-306-473A-116	Sequence 116, App
16	29	34.5	15	2	US-08-621-803-101	Sequence 101, App
17	29	34.5	15	2	US-08-485-445A-116	Sequence 116, App
18	29	34.5	15	3	US-09-119-263-116	Sequence 116, App
19	29	34.5	15	3	US-08-657-162-116	Sequence 116, App
20	29	34.5	15	3	US-09-224-180-116	Sequence 116, App
21	29	34.5	15	3	US-09-217-352-101	Sequence 101, App
22	29	34.5	15	5	PCR-US95-00498-116	Sequence 116, App
23	29	34.5	15	5	PCR-US95-00656-116	Sequence 116, App
24	28	33.3	11	4	US-09-236-160-1	Sequence 1, Appli
25	28	33.3	14	4	US-09-025-596-23	Sequence 23, Appl
26	28	33.3	16	1	US-07-985-691-17	Sequence 17, Appl
27	28	33.3	16	4	US-09-223-139-32	Sequence 32, Appl

NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03314  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/705-8410  
TELEFAX: 415/397-8338  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Binding-site  
LOCATION: 1  
OTHER INFORMATION: /note= "The residue at this position is bound to an Alkaloid by a sulfide."  
US-09-050-811-5

Query Match 36.9%; Score 31; DB 4; Length 10;  
Best Local Similarity 71.4%; Pred. No. 81;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKKQKDK 13  
Db 2 WKKKKKK 8

RESULT 3  
US-09-050-811-6  
Sequence 6, Application US/09050811  
Patent No. 6387700  
GENERAL INFORMATION:  
APPLICANT: Rice, Kevin G.  
APPLICANT: Wadhwa, Manpreet S.  
TITLE OF INVENTION: Peptides for Gene Delivery  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/050,811  
FILING DATE: 30-MAR-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03314  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/705-8410  
TELEFAX: 415/397-8338  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Binding-site  
LOCATION: 1  
OTHER INFORMATION: /note= "the residue at this position is bound to an Alkaloid by a Sulfide."

US-09-050-811-6

Query Match 36.9%; Score 31; DB 4; Length 15;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKKQKDK 13  
Db 2 WKKKKKK 8

RESULT 4  
US-09-050-811-11  
Sequence 11, Application US/09050811  
Patent No. 6387700  
GENERAL INFORMATION:  
APPLICANT: Rice, Kevin G.  
APPLICANT: Wadhwa, Manpreet S.  
TITLE OF INVENTION: Peptides for Gene Delivery  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MEDLEN & CARROLL, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/050,811  
FILING DATE: 30-MAR-1998  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UM-03314  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/705-8410  
TELEFAX: 415/397-8338  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-050-811-11

Query Match 36.9%; Score 31; DB 4; Length 15;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKKQKDK 13  
Db 2 WKKKKKK 8

RESULT 5  
US-08-621-259A-205  
Sequence 205, Application US/08621259A  
Patent No. 5858974  
GENERAL INFORMATION:  
APPLICANT: Little II, Roger G.  
APPLICANT: Lim, Edward  
APPLICANT: Padem, Mitchell B.  
TITLE OF INVENTION: Anti-Fungal Peptides  
NUMBER OF SEQUENCES: 252  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 11:31:08 ; Search time 67 Seconds  
(without alignments)  
37.782 Million cell updates/sec

Title: US-09-730-465-2\_COPY\_50\_65  
Perfect score: 84  
Sequence: 1 PLKEVLWKKOKDKVAE16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Total number of hits satisfying chosen parameters: 107419

```
Minimum DB seq length: 0
Maximum DB seq length: 16
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pcp:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:*
9: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:*
10: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:*
11: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:*
13: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:*
15: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pcp:*
18: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:*
19: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	36	42.9	16	9	US-09-853-233-15	Sequence 15, Appl
2	36	42.9	16	9	US-09-853-233-16	Sequence 16, Appl
3	34	40.5	15	9	US-09-853-233-17	Sequence 17, Appl
4	31	36.9	8	11	US-09-882-291-63	Sequence 63, Appl
5	31	36.9	9	11	US-09-882-291-61	Sequence 61, Appl
6	31	36.9	10	12	US-10-108-844-5	Sequence 5, Appl
7	31	36.9	11	11	US-09-882-291-55	Sequence 55, Appl
8	31	36.9	12	11	US-09-882-291-35	Sequence 35, Appl
9	31	36.9	12	11	US-09-882-291-50	Sequence 50, Appl
10	31	36.9	12	11	US-09-882-291-53	Sequence 53, Appl
11	31	36.9	13	11	US-09-882-291-56	Sequence 56, Appl
12	31	36.9	15	12	US-10-108-844-6	Sequence 6, Appl
13	31	36.9	15	12	US-10-108-844-11	Sequence 11, Appl
14	30	35.7	9	9	US-09-881-490-163	Sequence 163, Appl
15	29	34.5	15	9	US-09-765-537-101	Sequence 101, Appl

16	29	34.5	15	12	US-10-014-322A-90	Sequence 90, Appl
17	29	34.5	15	12	US-10-106-698-7620	Sequence 7620, Ap
18	28	33.3	12	11	US-09-882-291-38	Sequence 38, Appl
19	28	33.3	12	12	US-09-870-498A-1	Sequence 1, Appli
20	28	33.3	12	12	US-10-224-358-32	Sequence 32, Appl
21	28	33.3	14	12	US-10-224-356-34	Sequence 34, Appl
22	28	33.3	14	12	US-10-020-269-23	Sequence 23, Appl
23	28	33.3	12	12	US-10-014-322A-100	Sequence 100, App
24	28	33.3	16	12	US-10-225-567A-1191	Sequence 1191, Ap
25	27	32.1	9	12	US-10-356-829-33	Sequence 33, Appl
26	27	32.1	10	11	US-09-572-404B-663	Sequence 663, App
27	27	32.1	10	11	US-09-572-404B-665	Sequence 665, App
28	27	32.1	12	11	US-09-876-904A-247	Sequence 247, App
29	27	32.1	9	15	US-10-264-303-15	Sequence 15, Appl
30	26	31.0	9	10	US-09-071-838-74	Sequence 74, Appl
31	26	31.0	9	15	US-10-213-512-74	Sequence 74, Appl
32	26	31.0	10	10	US-09-984-056-85	Sequence 85, Appl
33	26	31.0	10	10	US-09-984-057-85	Sequence 85, Appl
34	26	31.0	10	11	US-09-572-404B-1947	Sequence 1947, Ap
35	26	31.0	10	12	US-10-105-233-85	Sequence 85, Appl
36	26	31.0	11	10	US-09-791-378-302	Sequence 302, App
37	26	31.0	11	15	US-10-116-391-10	Sequence 10, Appl
38	26	31.0	12	15	US-10-216-810A-5	Sequence 5, Appli
39	26	31.0	13	12	US-10-153-244-55	Sequence 55, Appl
40	26	31.0	13	12	US-10-153-244-119	Sequence 119, App
41	26	31.0	14	12	US-10-153-244-169	Sequence 169, App
42	26	31.0	13	12	US-10-153-244-218	Sequence 218, App
43	26	31.0	13	15	US-10-211-088-232	Sequence 232, App
44	26	31.0	14	9	US-09-766-397-8	Sequence 8, Appli
45	26	31.0	14	14	US-10-062-376-8	Sequence 8, Appli

## ALIGNMENTS

```

RESULT 1
US-09-853-253-15
/ Sequence 15, Application US/09853253
/ Patent No. US20020055156A1
/ GENERAL INFORMATION:
/ APPLICANT: JASPERS, STEPHEN
/ APPLICANT: SHEPPARD, PAUL
/ APPLICANT: DEISHER, THERESA
/ APPLICANT: BISHOP, PAUL
/ TITLE OF INVENTION: zsig33-like peptides
/ FILE REFERENCE: 00-30
/ CURRENT APPLICATION NUMBER: US/09/853,253
/ CURRENT FILING DATE: 2001-05-10
/ PRIOR APPLICATION NUMBER: 60/203,300
/ PRIOR FILING DATE: 2000-05-11
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 15
/ LENGTH: 16
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-853-253-15

```

```
Query Match      42.9%; Score 36; DB 9; Length 16;
Best Local Similarity 33.3%; Pred. No. 56;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
```

RESULT 2  
US-09-853-253-16  
; Sequence 16, Application US/09853253  
; Patent No. US2002005156A1  
; GENERAL INFORMATION:  
; APPLICANT: JASPERS, STEPHEN

```
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA
; APPLICANT: BISHOP, PAUL
; TITLE OF INVENTION: Zsig333-like Peptides
; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (16)...(16)
;
US-09-853-253-16

Query Match          42.9%; Score 36; DB 9; Length 16;
Best Local Similarity 33.3%; Pred. No. 56;
Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQKDKVAE 16
Db 2 LQDILWEEAKEAPAD 16

RESULT 3
US-09-853-253-17
; Sequence 17, Application US/09853253
; Patent No. US20020055156A1
; GENERAL INFORMATION:
; APPLICANT: JASPERS, STEPHEN
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA
; TITLE OF INVENTION: Zsig333-like Peptides
; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: AMIDATION
; LOCATION: (15)...(15)
;
US-09-853-253-17

Query Match          40.5%; Score 34; DB 9; Length 15;
Best Local Similarity 35.7%; Pred. No. 1e+02;
Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 LKEVLWKQKDKVA 15
Db 2 LQDILWEEAKEAP 15

RESULT 4
US-09-882-291-63
; Sequence 63, Application US/09882291
; Publication No. US2003004072A1
; GENERAL INFORMATION:
; APPLICANT: Zealand Pharmaceuticals A/S
; TITLE OF INVENTION: No. US2003004072A1el Peptide Conjugates
; FILE REFERENCE: 007-2001
; CURRENT APPLICATION NUMBER: US/09/882,291
; CURRENT FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequen
US-09-882-291-61

Query Match          36.9%; Score 31; DB 11; Length 8;
Best Local Similarity 71.4%; Pred. No. 5.2e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKQKQDK 13
Db 2 WKKKKKK 8

RESULT 5
US-09-882-291-61
; Sequence 61, Application US/09882291
; Publication No. US2003004072A1
; GENERAL INFORMATION:
; APPLICANT: Zealand Pharmaceuticals A/S
; TITLE OF INVENTION: No. US2003004072A1el Peptide Conjugates
; FILE REFERENCE: 007-2001
; CURRENT APPLICATION NUMBER: US/09/882,291
; CURRENT FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic peptide sequen
US-09-882-291-61

Query Match          36.9%; Score 31; DB 11; Length 8;
Best Local Similarity 71.4%; Pred. No. 5.2e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WKQKQDK 13
Db 2 WKKKKKK 8

RESULT 6
US-10-108-844-5
; Sequence 5, Application US/10108844
; Publication No. US20030170894A1
; GENERAL INFORMATION:
; APPLICANT: Rice, Kevin G.
; TITLE OF INVENTION: Peptides for Gene Delivery
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/108,844
; FILING DATE: 28-Mar-2002
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:34:43 ; Search time 102.059 Seconds  
(without alignments)  
539.670 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVCLL.....MHEALHNYTKSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*  
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1856	100.0	347	13	CD2 binding LFA-3-
2	1856	100.0	347	21	Human transmembran
3	1856	100.0	347	22	Human LFA3TIP fusi
4	1856	100.0	347	23	Human lymphocyte f
5	1856	100.0	347	23	Human LFA3TIP poly
6	1856	100.0	347	23	Human LFA3TIP fusion pr
7	1856	100.0	347	24	Murine LFA-3. Mus
8	1845	99.4	351	23	ABP58178
9	1834	98.8	347	14	Human LFA-3/IgG fu

10	1366	73.6	256	24	AA016017
11	1265	68.2	458	22	AA97596
12	1265	68.2	458	23	ABP52449
13	1263.5	68.1	432	23	ABP52450
14	1263.5	68.1	458	22	AA97594
15	1263.5	68.1	458	23	ABP52447
16	1262	68.0	744	21	AA94408
17	1255.5	67.6	633	21	AA84965
18	1255	67.6	477	20	AAW90207
19	1251.5	67.4	388	19	AAW73514
20	1251.5	67.4	388	21	AAW54064
21	1250.5	67.4	388	19	AAW73513
22	1250.5	67.4	388	21	AAW54063
23	1249.5	67.3	497	21	AA97172
24	1249.5	67.3	525	21	AA97171
25	1249	67.3	859	20	AAW70796
26	1249	67.3	859	21	AA922184
27	1249	67.3	1158	21	AA92205
28	1249	67.3	1168	21	AA92204
29	1248	67.2	455	22	AA97595
30	1248	67.2	455	23	ABP52448
31	1247	67.2	414	23	AA666031
32	1247	67.2	437	24	ABJ37104
33	1247	67.2	446	15	AA858753
34	1247	67.2	446	20	AA923986
35	1247	67.2	446	20	AA97595
36	1247	67.2	446	20	AA96743
37	1247	67.2	446	24	ABU08540
38	1247	67.2	451	23	AAU081014
39	1247	67.2	480	24	AAU016238
40	1247	67.2	480	24	AAU016239
41	1247	67.2	698	23	AAU016239
42	1247	67.2	698	24	AAU016237
43	1246	67.1	480	20	AAW90206
44	1246	67.1	480	24	ABU07263
45	1246	67.1	1367	19	AAW70542

#### ALIGNMENTS

RESULT 1  
AA927163  
ID AAR27163 standard; Protein; 347 AA.  
XX AAR27163;  
AC  
XX  
XX  
DT 25-MAR-2003 (updated)  
DT 20-MAY-1998 (first entry)  
XX  
DE CD2 binding LFA-3-Ig fusion protein.  
XX  
XX lymphocyte associated antigen-3; T-lymphocyte accessory molecule;  
KW deletion mutant; CD2 binding site; immunomodulator; immunoglobulin;  
KW preLFA3TIP; LFA-3(92)IgG; PSAB152; ss.  
XX Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT peptide 1..28  
FT /label= LFA-3\_signal  
FT region 29..120  
FT /label= LFA-3\_aminoacids\_1-92  
FT region 121..130  
FT /label= IgG1\_hinge  
FT domain 131..239  
FT /label= IgG1\_CH2  
FT domain 240..347  
FT /label= IgG1\_CH3  
XX  
XX EP503648-A1.  
XX  
XX 16-SEP-1992.

Murine CD2-binding  
Ft1 receptor fusi  
Modified Ft1 rece  
Ft1D2.Ft1D3.FGDE  
Ft1 receptor fusi  
Modified Ft1 rece  
Human VCAM-1/IgG1-  
Human VCAM-1/IgG1-  
Amino acid sequenc  
hB7.2Fc soluble fu  
Human TGFbetaRII:F  
Amino acid sequenc  
Rabbit TGFbetaRII:  
Human FGF-RI Extra  
Human FGF-RI Extra  
Human gp130-Fc-His  
Human gp130-Fc-His  
Fusion polypeptide  
Fusion polypeptide  
Ft1 receptor fusi  
Modified Ft1 rece  
Human RYK fragment  
Concatermic immun  
VCAM 2D-IgG. Homo  
VCAM 2D-IgG, a sol  
VCAM 2D-IgG protei  
A VCAM 2D-IgG1 fus  
Human VCAM-1/IgG f  
B7-related protein  
B7-related protein  
B7-related protein  
B7-related protein  
hB7.1Fc soluble fu  
Human expressed pr  
Integrin alpha-2 c



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:45:48 ; Search time 42.5245 Seconds  
(without alignments)  
784.736 Million cell updates/sec

Title: US-09-730-465-8  
Perfect score: 1856  
Sequence: 1 MVASDAGRALGVLSVCLL.....MHEALNHVYTKSLSLSPGK 347

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76: \*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1233	66.4	255	4 S31866	Ig gamma-1 chain C
2	1233	66.4	330	1 GHU	Ig gamma-1 chain C
3	1228.5	66.2	374	2 S89339	Ig heavy chain V r
4	1180	63.6	234	2 P70207	Ig gamma chain C r
5	1146	61.7	377	2 A23511	Ig gamma-3 chain C
6	1145.5	61.7	326	1 G2HU	Ig gamma-2 chain C
7	1144	61.6	377	2 A50764	Ig gamma-3 chain C
8	1135	61.2	327	1 G4HU	Ig gamma-4 chain C
9	1121	60.4	289	1 G3HWI	Ig gamma-3 heavy C
10	926.5	49.9	323	1 GHRB	Ig gamma chain C r
11	913.5	49.2	328	2 I47160	Ig gamma 2b chain
12	909.5	49.0	328	2 I47159	Ig gamma 2a chain
13	906.5	48.8	277	2 I47162	Ig gamma 4 chain c
14	894.5	48.2	328	2 I47158	Ig gamma 1 chain c
15	892	48.1	329	1 G2GP	Ig gamma-2 chain C
16	886.5	47.8	328	2 I47161	Ig gamma 3 chain c
17	857.5	46.2	470	2 S22080	Ig heavy chain pre
18	846	45.6	308	2 C30554	Ig heavy chain C r
19	846	45.6	472	2 S31459	Ig gamma-1 chain -
20	845.5	45.6	329	1 G3MSC	Ig gamma-3 chain C
21	838	45.2	333	2 PS0018	Ig gamma-2b chain
22	834.5	45.0	398	1 G3MSM	Ig gamma-3 chain C
23	829.5	44.7	444	2 P24436	monoclonal antibod
24	819	44.1	326	2 PS0017	Ig gamma-1 chain C
25	818.5	44.1	324	1 G1MS	Ig gamma-1 chain C
26	813.5	43.8	393	1 G3MSM	Ig gamma-1 chain C
27	812	43.8	329	2 S00847	Ig gamma-2c chain
28	811.5	43.7	330	1 G2MSA	Ig gamma-2a chain
29	811.5	43.7	469	2 S37483	Ig gamma-2a chain

30	806.5	43.5	399	1 G2MSAM	Ig gamma-2a chain
31	804	43.3	335	1 G2MSAB	Ig gamma-2a chain
32	796.5	42.9	446	2 S40295	Ig gamma-2a chain
33	787.5	42.4	322	2 PS0019	Ig gamma-2a chain
34	783.5	42.2	474	1 G2MS11	Ig gamma-2b chain
35	778.5	41.9	405	1 G2MSBM	Ig gamma-2b chain
36	767.5	41.4	475	2 S01321	Ig gamma-2b chain
37	764	41.2	327	2 S06611	Ig gamma-2 chain C
38	707	38.1	180	2 I46732	Ig gamma heavy cha
39	630	33.9	250	2 A28564	lymphocyte functio
40	577.5	31.1	249	2 S69340	Ig heavy chain VHI
41	574.5	31.0	218	2 A36040	Ig heavy chain V-I
42	571	30.8	252	2 S14236	Ig gamma-1 chain C
43	395.5	21.3	572	2 B46529	Ig x heavy chain (
44	368	19.8	388	1 EHMS	Ig epsilon chain C
45	366.5	19.7	429	1 EHRT	Ig epsilon chain C

ALIGNMENTS

RESULT 1

S31866  
Ig gamma-1 chain C region - synthetic  
C:Species: synthetic  
A:Note: Homo sapiens (man) gene engineered and expressed in Escherichia coli  
C:Date: 06-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 19-May-2000  
C:Accession: S31866  
R:Filpula, D.  
submitted to the EMBL Data Library, February 1993  
A:Description: Screening method for protein-protein interactions of cloned gene pro  
A:Reference number: S31866  
A:Accession: S31866  
A:Molecule type: mRNA  
A:Residues: 1-255 <FIL>  
A:Cross-references: EMBL:X70421; NID:g33068; PIDN:CAA49866.1; PID:g33069  
C:Keywords: immunoglobulin  
F:1-22/Region: Escherichia coli outer membrane protein A precursor  
F:23-255/Region: human Ig gamma-1 chain C region

Query Match	66.4%	Score	1233	DB	4	Length	255
Best Local Similarity	100.0%	Pred. No.	1.5e-80				
Matches	227	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	121	DKHTCPCCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD	180				
Db	29	DKHTCPCCPAPPELLGGPSVFLFPPPKKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYD	88				
QY	181	GVEVHNAKTPEEQYNSTYRVVSVLTVLDQWLNGKEYKCKVSNKALPAPIETISKAK	240				
Db	89	GVEVHNAKTPEEQYNSTYRVVSVLTVLDQWLNGKEYKCKVSNKALPAPIETISKAK	148				
QY	241	GQREPOVYITLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	300				
Db	149	GQREPOVYITLPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTPPVLD	208				
QY	301	DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK	347				
Db	209	DGSFFLYSKLTVDKSRWQQGNVSCFVSVMHEALHNHYTQKSLSLSPGK	255				

RESULT 2

GHU  
Ig gamma-1 chain C region - human  
C:Species: Homo sapiens (man)  
C:Date: 31-Jan-1981 #sequence\_revision 18-Aug-1982 #text\_change 16-Jul-1999  
C:Accession: A93433; S36861; S33887; B90563; A90564; B91668; A91723; A02146  
R:Ellison, J.W.; Berson, B.J.; Hood, L.E.  
Nucleic Acids Res. 10, 4071-4079, 1982.  
A:Title: The nucleotide sequence of a human immunoglobulin C-gamma1 gene.  
A:Reference number: A93433; MUID:82274238; PMID:6287432  
A:Accession: A93433  
A:Molecule type: DNA



hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate in a C; Superfamily: immunoglobulin C region; immunoglobulin homology  
C; Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin  
F; 20-85/Domain: immunoglobulin homology <IM1>  
F; 137-206/Domain: immunoglobulin homology <IM2>  
F; 243-310/Domain: immunoglobulin homology <IM3>  
F; 27-63,144-204,250-308/Disulfide bonds: #status experimental  
F; 103/Disulfide bonds: interchain (to light chain) #status experimental  
F; 109,112/Disulfide bonds: interchain (to heavy chain) #status experimental  
F; 180/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 66.4%; Score 1233; DB 1; Length 330;  
Best Local Similarity 100.0%; Pred. No. 2.1e-80;  
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DKTHTCPCPAPELLGGPSVFLFPPKPKDRLMSRTPEVTCVVVDVSHEDPEVKFNWYVD 180  
Db 104 DKTHTCPCPAPELLGGPSVFLFPPKPKDRLMSRTPEVTCVVVDVSHEDPEVKFNWYVD 163  
QY 181 GVEVHNAKTPREEQNSYRYVSVLTVLHODWLNGKEYCKVSNKALPAPTEKTIISKAK 240  
Db 164 GVEVHNAKTPREEQNSYRYVSVLTVLHODWLNGKEYCKVSNKALPAPTEKTIISKAK 223  
QY 241 GQPREPQVITLPPSRDELAKNOVSLTCLVKGYFSDIAVEWESNGQPENNYKTTTPPVLD 300  
Db 224 GQPREPQVITLPPSRDELAKNOVSLTCLVKGYFSDIAVEWESNGQPENNYKTTTPPVLD 283  
QY 301 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 347  
Db 284 DGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK 330

RESULT 3  
S69339  
Ig heavy chain V region precursor - human  
C; Species: Homo sapiens (man)  
C; Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 01-Dec-2000  
C; Accession: S69339; S72664  
R; Khamlichi, A.A.; Autcourtier, P.; Preud'homme, J.L.; Cogne, M.  
Eur. J. Biochem. 229, 54-60, 1995  
A; Title: Structure of abnormal heavy chains in human heavy-chain-deposition disease  
A; Reference number: S69339; MUID:95262687; PMID:7744049  
A; Accession: S69339  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-374 <KHA>  
A; Cross-references: EMBL:X81695  
R; Khamlichi, A.A.  
submitted to the EMBL Data Library, September 1994  
A; Reference number: S72664  
A; Accession: S72664  
A; Status: preliminary  
A; Molecule type: mRNA  
A; Residues: 1-140, 'C', 142-374 <KH2>  
A; Cross-references: EMBL:X81695  
C; Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 66.2%; Score 1228.5; DB 2; Length 374;  
Best Local Similarity 79.3%; Pred. No. 5.2e-80;  
Matches 234; Conservative 10; Mismatches 22; Indels 29; Gaps 1;

QY 82 YLDTVSGSLTIYNLTGSDDEYEMESPNITDMDKFFL----- 118  
Db 80 YPSLFRFLRTITKDTGRKNQVLTINVDPADATYATYCGYSVGYGOGYRPHSWGOGTLVIT 139  
QY 119 -----YVDKTHTCPCPAPELLGGPSVFLFPPKPKDRLMSRTPEVTCVVVDVSHEDPE 172  
Db 140 VSSEPKSCDKTHTCPCPAPELLGGPSVFLFPPKPKDRLMSRTPEVTCVVVDVSHEDPE 199  
QY 173 VKFNWYVDGVEVHNAKTPREEQNSYRYVSVLTVLHODWLNGKEYCKVSNKALPAPTE 232  
Db 200 VKFNWYVDGVEVHNAKTPREEQNSYRYVSVLTVLHODWLNGKEYCKVSNKALPAPTE 259

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:35:28 ; Search time 22.9632 Seconds  
(without alignments)  
710.626 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....MHEALHNHYTKSLSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1293	66.4	330	1	GCL_HUMAN
2	1145.5	61.7	326	1	GCL_HUMAN
3	1135	61.2	327	1	GCL_HUMAN
4	1126	60.7	290	1	GCL_HUMAN
5	926.5	49.9	323	1	GCL_HUMAN
6	892	48.1	329	1	GCL_HUMAN
7	845.5	45.6	329	1	GCL_HUMAN
8	838	45.2	333	1	GCL_HUMAN
9	834.5	45.0	398	1	GCL_HUMAN
10	819	44.1	326	1	GCL_HUMAN
11	818.5	44.1	324	1	GCL_HUMAN
12	813.5	43.8	393	1	GCL_HUMAN
13	812	43.8	329	1	GCL_HUMAN
14	811.5	43.7	330	1	GCL_HUMAN
15	806.5	43.5	399	1	GCL_HUMAN
16	804	43.3	335	1	GCL_HUMAN
17	787.5	42.4	322	1	GCL_HUMAN
18	783.5	42.2	336	1	GCL_HUMAN
19	778.5	41.9	405	1	GCL_HUMAN
20	630	33.9	420	1	GCL_HUMAN
21	368	19.8	421	1	GCL_HUMAN
22	366.5	19.7	429	1	GCL_HUMAN
23	360.5	19.4	428	1	GCL_HUMAN
24	358.5	19.3	454	1	GCL_HUMAN
25	357	19.2	455	1	GCL_HUMAN
26	356.5	19.2	391	1	GCL_HUMAN
27	351	18.9	458	1	GCL_HUMAN
28	347	18.7	476	1	GCL_HUMAN
29	341	18.4	479	1	GCL_HUMAN
30	335	18.0	457	1	GCL_HUMAN
31	332	17.9	450	1	GCL_HUMAN
32	332	17.9	454	1	GCL_HUMAN
33	302.5	16.3	438	1	GCL_HUMAN

34	298.5	16.1	438	1	HVCS_HETFR	P23087 heterodontu
35	295.5	15.9	299	1	ALC_RABIT	P01879 oryctolagus
36	291	15.7	446	1	MUC_CHICK	P01875 gallus gall
37	289.5	15.6	461	1	HVCM_HETFR	P23088 heterodontu
38	282.5	15.2	393	1	HVC3_HETFR	P23086 heterodontu
39	277.5	15.0	370	1	HVC1_HETFR	P23084 heterodontu
40	272	14.7	353	1	ALC1_HUMAN	P01876 homo sapien
41	271.5	14.6	340	1	ALC2_HUMAN	P01877 homo sapien
42	268	14.4	353	1	ALC1_GORGO	P20758 gorilla gor
43	257.5	13.9	481	1	MUCM_ICTPU	P23735 ictalurus p
44	245	13.2	344	1	ALC_MOUSE	P01878 mus musculu
45	205	11.0	513	1	SHSL_MOUSE	P97797 m protein-t

## ALIGNMENTS

RESULT 1	GCL_HUMAN	STANDARD;	PRT;	330 AA.
ID	GCL_HUMAN			
AC	P01857;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DE	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Ig gamma-1 chain C region.			
GN	IGHG1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=82274238; PubMed=6287432;			
RA	Ellison J.W., Berson B.J., Hood L.E.;			
RT	"The nucleotide sequence of a human immunoglobulin C gamma1 gene."			
RL	Nucleic Acids Res. 10:4071-4079(1982).			
RN	[2]			
RP	SEQUENCE OF 1-135 (MYELOMA PROTEIN EU).			
RX	MEDLINE=71064024; PubMed=5489771;			
RA	Cunningham B.A., Rutishauser U., Gall W.E., Gottlieb P.D.,			
RA	Waxdal M.J., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. VII. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H1-H4."			
RN	Biochemistry 9:3161-3170(1970).			
RP	SEQUENCE OF 136-329 (EU).			
RX	MEDLINE=71064025; PubMed=5530842;			
RA	Rutishauser U., Cunningham B.A., Bennett C., Konigsberg W.H.,			
RA	Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. 8. Amino			
RL	acid sequence of heavy-chain cyanogen bromide fragments H5-H7."			
RN	Biochemistry 9:3171-3181(1970).			
RP	SEQUENCE (MYELOMA PROTEIN NIE).			
RX	MEDLINE=77070269; PubMed=826475;			
RA	Ponstingl H., Hilschmann N.;			
RT	"The rule of antibody structure. The primary structure of a			
RL	monoclonal IgG1 immunoglobulin (myeloma protein NIE). III. The			
RN	chymotryptic peptides of the H-chain, alignment of the tryptic			
RP	peptides and discussion of the complete structure."			
RX	Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).			
RL	[5]			
RP	SEQUENCE (MYELOMA PROTEIN KOL), AND DISULFIDE BONDS.			
RX	MEDLINE=83289131; PubMed=6884994;			
RA	Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;			
RT	"Three-dimensional structure determination of antibodies. Primary			
RL	structure of crystallized monoclonal immunoglobulin IgG1 KOL, I."			
RN	Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).			
RP	[6]			
RX	DISULFIDE BONDS.			
RX	MEDLINE=71064027; PubMed=4923144;			
RA	Gall W.E., Edelman G.M.;			
RT	"The covalent structure of a human gamma G-immunoglobulin. X.			

Intrachain disulfide bonds.";  
 Biochemistry 9:3188-3196(1970).  
 [7]  
 DISULFIDE BONDS.  
 RA MEDLINE=77070267; PubMed=1002129;  
 RX Dreker L., Schwarz J., Reichel W., Hilschmann N.;  
 RL "Rule of antibody structure. The primary structure of a monoclonal  
 RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and  
 RT characterization of the protein, the L- and H-chains, the  
 RT cyanogen bromide cleavage products, and the disulfide bridges.";  
 RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).  
 [8]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=81208100; PubMed=7236608;  
 RA Deisenhofer J.;  
 RL "Crystallographic refinement and atomic models of a human Fc fragment  
 RT aureus at 2.9- and 2.8-A resolution.";  
 RL Biochemistry 20:2361-2370(1981).  
 CC GLM(1) MARKERS: NIE HAS THE GLM(17) ALLOTYPIC MARKER, 97-K, & THE  
 CC GLM(1) MARKERS, 239-D & 241-L. KOL & EU SEQUENCES HAVE THE GLM(3)  
 CC MARKER & THE GLM (NON-1) MARKERS.  
 CC -I- MISCELLANEOUS: NIE ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC 35,116,198,269 & 272.  
 CC -I- MISCELLANEOUS: EU ALSO DIFFERS IN THE AMIDATION STATES OF RESIDUES  
 CC 155, 166, 177, 195, 198, 269, AND 272 AND IN THE ORDER OF RESIDUES  
 CC 268-272.  
 CC -I- MISCELLANEOUS: KOL ALSO DIFFERS IN THE AMIDATION STATES OF  
 CC RESIDUES 198,267&272.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; J00228; AAC82527.1; ALT\_INIT.  
 DR PIR; A93433; GHU.  
 DR PDB; 1FC1; 15-JUL-92.  
 DR PDB; 1FC2; 15-JUL-92.  
 DR PDB; 1A77; 12-NOV-97.  
 DR PDB; 1D5B; 09-FEB-00.  
 DR PDB; 1D5I; 09-FEB-00.  
 DR PDB; 1D6V; 04-OCT-00.  
 DR PDB; 1DN2; 17-MAY-00.  
 DR PDB; 1E4K; 06-JUN-01.  
 DR PDB; 1FCC; 20-JUL-95.  
 DR PDB; 1H2H; 12-JUN-02.  
 DR PDB; 1I7Z; 08-AUG-01.  
 DR PDB; 1IIS; 16-MAY-01.  
 DR PDB; 1IIX; 16-MAY-01.  
 DR PDB; 1L6X; 10-APR-02.  
 DR PDB; 2RGS; 12-NOV-97.  
 DR Genew; HGNC:5525; IGHL1.  
 DR MIM; 147100; -.  
 DR GO; GO:0005624; C:membrane fraction; NAS.  
 DR GO; GO:0003823; F:antigen binding activity; TAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003597; Ig-CL.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR Pfam; PF00047; Ig; 3.  
 DR SMART; SM00407; IGCL1; 2.  
 DR PROSITE; PS00835; IG\_LIKE; 3.  
 DR PROSITE; PS00290; IG\_MHC; 2.  
 DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;  
 3D-structure.  
 KW NON\_TER 1 1  
 KW DOMAIN 1 98 CH1.  
 FT FT 110 HINGE.  
 FT DOMAIN 99 110  
 FT DOMAIN 111 223 CH2.

CH3.  
 INTERCHAIN (WITH LIGHT CHAIN).  
 INTERCHAIN (WITH HEAVY CHAIN).  
 INTERCHAIN (WITH HEAVY CHAIN).  
 N-LINKED (GLCNAC...).  
 REMOVED POST-TRANSLATIONALLY.  
 K -> R (IN GLM(3) MARKER).  
 /FTIG-VAR\_003886.  
 D -> E (IN GLM(NON-1) MARKER).  
 /FTIG-VAR\_003887.  
 L -> M (IN GLM(NON-1) MARKER).  
 /FTIG-VAR\_003888.  
 224 330  
 27 83  
 103 103  
 109 109  
 112 112  
 144 204  
 250 308  
 180 180  
 330 330  
 97 97  
 239 239  
 241 241  
 122 126  
 130 134  
 136 137  
 141 147  
 157 162  
 163 164  
 185 166  
 188 171  
 176 179  
 180 181  
 182 190  
 193 197  
 198 199  
 202 207  
 209 210  
 215 219  
 227 227  
 230 234  
 238 242  
 245 256  
 260 265  
 270 270  
 274 276  
 280 281  
 283 284  
 287 296  
 301 301  
 302 303  
 305 312  
 313 314  
 316 317  
 320 325  
 330 AA; 36106 MW; 3770EEL06C2FA33D CRC64;  
 Query Match 66.4%; Score 1233; DB 1; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-85;  
 Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 121 DKTHCTPCPCAPPELLGSPVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYD 180  
 DB 104 DKTHCTPCPCAPPELLGSPVFLFPKPKDITLMISRTPEVTCVVVDVSHEDPEVKFNWYD 163  
 QY 181 GVEVHNATKPREQYNSTYRVVSVLTVTHQDNLGKEYCKVSNKALPAPTEKTSKAK 240  
 DB 164 GVEVHNATKPREQYNSTYRVVSVLTVTHQDNLGKEYCKVSNKALPAPTEKTSKAK 223  
 QY 241 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 300  
 DB 224 GQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD 283  
 QY 301 DGSFFLYSKLTVDKSRWQQGNVFCVMHSEALHNHYTQKSLSLSPGK 347  
 DB 284 DGSFFLYSKLTVDKSRWQQGNVFCVMHSEALHNHYTQKSLSLSPGK 330  
 RESULT 2  
 GC2\_HUMAN  
 ID GC2\_HUMAN STANDARD; PRT; 326 AA.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:45:18 ; Search time 102.909 Seconds  
(without alignments)  
870.128 Million cell updates/sec

Title: US-09-730-465-8

Sequence: 1 MVAGSDAGRALGVLSVVCILL.....MHEALHNYTKSLSPCK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mbc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1233	66.4	471	4	Q8TC77	Q8tc77 homo sapien
2	1233	66.4	701	4	Q96P08	Q96pq8 homo sapien
3	1142	61.5	521	4	Q8N4Y9	Q8n4y9 homo sapien
4	1135	61.2	473	4	Q8TC63	Q8tc63 homo sapien
5	1131	60.9	509	4	Q8NF17	Q8nf17 homo sapien
6	909	49.0	337	6	Q95M34	Q95m34 equus caball
7	831.5	44.8	469	11	Q8R3V9	Q8r3v9 mus musculus
8	830.5	44.7	463	11	Q991C4	Q991c4 mus musculus
9	826.5	44.5	437	11	Q9RIA4	Q9r1a4 mus musculus
10	811	43.7	473	11	Q9D8L4	Q9d8l4 mus musculus
11	807.5	43.5	468	11	Q991J1	Q991j1 mus musculus
12	807.5	43.5	473	11	Q991J5	Q991j5 mus musculus
13	784.5	42.3	473	11	Q91Z05	Q91z05 mus musculus
14	784.5	42.3	474	11	Q8R3H6	Q8r3h6 mus musculus
15	632.5	34.1	240	4	Q9BRW0	Q9brw0 homo sapien
16	476	25.6	134	4	Q14748	Q14748 homo sapien

17	436	23.5	119	4	Q16393	Q16393 homo sapien
18	358.5	19.3	375	4	Q9BSZ1	Q9bsz1 homo sapien
19	358.5	19.3	597	4	Q9BU10	Q9bu10 homo sapien
20	358.5	19.3	597	4	Q9BQB8	Q9bqb8 homo sapien
21	358.5	19.3	597	4	Q96BB9	Q96bb9 homo sapien
22	348.5	18.8	588	4	Q8WUX4	Q8wux4 homo sapien
23	348.5	18.8	613	4	Q96EY0	Q96ey0 homo sapien
24	348.5	18.8	613	4	Q8WUK1	Q8wuk1 homo sapien
25	348.5	18.8	614	4	Q96GA6	Q96ga6 homo sapien
26	348.5	18.8	618	4	Q96AA6	Q96aa6 homo sapien
27	347	18.7	613	11	Q8VCX7	Q8vcx7 mus musculus
28	329.5	17.8	227	6	Q28754	Q28754 ovis sp. lf
29	329.5	17.8	253	6	Q28753	Q28753 ovis sp. lf
30	319.5	17.2	159	6	Q28752	Q28752 ovis sp. lf
31	280.5	15.1	684	13	Q90544	Q90544 ginglymosto
32	272	14.7	384	4	Q9UP60	Q9up60 homo sapien
33	272	14.7	493	4	Q8NCL6	Q8ncl6 homo sapien
34	272	14.7	494	4	Q96K68	Q96k68 homo sapien
35	272	14.7	496	4	Q96KX8	Q96kx8 homo sapien
36	271.5	14.6	416	4	Q9NEP6	Q9nep6 homo sapien
37	269	14.5	496	4	Q96DK0	Q96dk0 homo sapien
38	269	14.5	499	4	Q8NSK4	Q8nsk4 homo sapien
39	267	14.4	497	4	Q8WY24	Q8wy24 homo sapien
40	267	14.4	500	4	Q9BRV0	Q9brv0 homo sapien
41	261	14.1	486	11	Q91Z07	Q91z07 mus musculus
42	261	14.1	487	11	Q99KA4	Q99ka4 mus musculus
43	260	14.0	426	11	Q9DCD9	Q9dcd9 mus musculus
44	253.5	13.7	481	11	Q91WT3	Q91wt3 mus musculus
45	253.5	13.7	481	11	Q91WT1	Q91wt1 mus musculus

#### ALIGNMENTS

#### RESULT 1

Q8TC77 Q8TC77 PRELIMINARY; PRT; 471 AA.  
AC Q8TC77;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Spleen;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC024289; AAh24289.1; -.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; Ig; 4.  
DR SMART; SM00406; IGV; 1.  
DR PROSITE; PS00835; IG\_LIKE; 4.  
DR PROSITE; PS00290; IG\_MHC; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 471 AA; 388F7F4CF588660E CRC64;

Query Match 66.4%; Score 1233; DB 4; Length 471;

Best Local Similarity 100.0%; Pred. No. 1.7e-100;

Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DKHTCCPCPAPELLGGPSVFLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180

Db 245 DKHTCCPCPAPELLGGPSVFLFPPKPKDLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 304

QY 181 GVEVHNKATPREEQYNSTYRVVSVLTVQLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240

Db 305 GVEVHNKATPREEQYNSTYRVVSVLTVQLHQQDWLNGKEYKCKVSNKALPAPIEKTISKAK 364

```

241 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 300
|||||
365 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 424
|||||
301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 347
|||||
425 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 471
|||||

RESULT 2
Q96PQ8 PRELIMINARY; PRT; 701 AA.
MEDLINE=21477448; PubMed=11593034;
Hu Z., Garen A.;
"Targeting tissue factor on tumor vascular endothelial cells and tumor
cells for immunotherapy in mouse models of prostatic cancer.";
Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
EMBL; AF272774; AAK58686.1; -.
HSSP; P00761; IANI.
InterPro; IPR000152; Asx_hydroxyl.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR000742; EGF_2.
InterPro; IPR001681; EGF_Ca.
InterPro; IPR001438; EGF_II.
InterPro; IPR006209; EGF-like.
InterPro; IPR002383; GLA_blood.
InterPro; IPR007110; Ig-like.
InterPro; IPR003597; Ig-cl.
InterPro; IPR003006; Ig-MHC.
InterPro; IPR001254; Ser_protease_Try.
InterPro; IPR000294; VitK_dep_GLA.
Pfam; PF00008; EGF; 2.
Pfam; PF00594; gla; 1.
Pfam; PF00047; Ig; 2.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00010; EGFLOOD.
PRINTS; PR00001; GLABLOOD.
SMART; SM00179; EGF_CA; 1.
SMART; SM00069; GLA; 1.
SMART; SM00407; Igcl; 1.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00010; ASX_HYDROXYL; 1.
PROSITE; PS00022; EGF_1; 1.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01187; EGF_CA; 1.
PROSITE; PS00011; GLU_CARBOXYLATION; 1.
PROSITE; PS50835; IG-LIKE; 2.
PROSITE; PS00290; IG-MHC; 1.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Hydrolase; Protease; Serine protease.
SEQUENCE 701 AA; 77826 MW; 94AC6CEB42CC992F CRC64;

```

```

Query Match 66.4%; Score 1233; DB 4; Length 701;
Best Local Similarity 100.0%; Pred. No. 2.9e-100;
Matches 227; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180

```

```

Db 475 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 534
|||||
QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
|||||
Db 535 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 594
|||||
QY 241 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 300
|||||
Db 595 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 654
|||||
QY 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 347
|||||
Db 655 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 701
|||||

RESULT 3
Q8N4Y9 PRELIMINARY; PRT; 521 AA.
ID Q8N4Y9
AC Q8N4Y9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Primary B-Cells from Tonsils;
RA Strausberg R.;
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
EMBL; BC033178; AAH33178.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig-MHC.
DR InterPro; IPR003596; Ig-V.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; Igcl; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00835; IG-LIKE; 4.
DR PROSITE; PS00290; IG-MHC; 2.
KW Hypothetical protein.
SQ SEQUENCE 521 AA; 57156 MW; 2AC7D22E72D6CAA2 CRC64;

Query Match 61.5%; Score 1142; DB 4; Length 521;
Best Local Similarity 92.1%; Pred. No. 2.2e-92;
Matches 209; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 121 DKHTCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 180
|||||
Db 295 DTTPPCPCPAPPELLGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVD 354
|||||
QY 181 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 240
|||||
Db 355 GVEVHNAKTPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAK 414
|||||
QY 241 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 300
|||||
Db 415 GQPREQVYTLPPSRDELTKNOVSLTCLVKGFPESDIATVWESNGOPENNYKTTTPPVLD 474
|||||
QY 301 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 347
|||||
Db 475 DGSFFLYSKLTVDKSRWQGNVFCSCVMHEALHNHYTKLSLSLSPGK 521
|||||

RESULT 4
Q8TC63 PRELIMINARY; PRT; 473 AA.
ID Q8TC63
AC Q8TC63;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2003, 16:47:23 ; Search time 34.8701 Seconds  
(without alignments)  
421.045 Million cell updates/sec

Title: US-09-730-465-8

Perfect score: 1856

Sequence: 1 MVAGSDAGRALGVLSVVCLL.....MHEALHNHYTKSLSPGK 347

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

#### Database :

Issued Patents.AA.\*  
1: /cgn2.6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/iaa/PCUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1856	100.0	347	1	US-07-940-861-43
2	1856	100.0	347	1	US-08-459-512-43
3	1856	100.0	347	1	US-08-459-657-43
4	1856	100.0	347	2	US-08-460-132-43
5	1856	100.0	347	3	US-08-466-465-8
6	1856	100.0	347	5	PCT-US92-02050-43
7	1249	67.3	859	4	US-09-313-942-7
8	1249	67.3	1158	4	US-09-313-942-26
9	1249	67.3	1168	4	US-09-313-942-24
10	1247	67.2	446	4	US-09-157-452B-12
11	1242	66.8	388	3	US-09-131-247-16
12	1239.5	66.8	482	4	US-09-189-129-2
13	1239	66.8	442	5	PCT-US96-10043-9
14	1239	66.8	704	4	US-09-590-656-2
15	1239	66.8	704	4	US-09-733-764-2
16	1238.5	66.7	424	5	PCT-US95-03866-12
17	1238.5	66.7	424	5	PCT-US95-03866-14
18	1238.5	66.7	437	5	PCT-US96-10043-11
19	1238.5	66.7	691	4	US-09-313-942-20
20	1238.5	66.7	694	4	US-09-313-942-22
21	1238	66.7	680	3	US-08-227-496C-15
22	1237	66.6	229	4	US-09-122-144-2
23	1236	66.6	592	4	US-09-313-942-8
24	1235.5	66.6	488	3	US-08-776-511-2
25	1234	66.5	552	1	US-08-243-010-6
26	1233.5	66.5	594	4	US-09-313-942-18
27	1233.5	66.5	793	4	US-09-313-942-32

28 1233 66.4 232 2 US-08-595-043A-50  
29 1233 66.4 235 3 US-09-131-247-6  
30 1233 66.4 331 3 US-09-178-869-2  
31 1233 66.4 331 4 US-09-761-413-2  
32 1233 66.4 360 4 US-09-180-100-11  
33 1233 66.4 371 1 US-08-236-311-7  
34 1233 66.4 371 3 US-08-457-918-7  
35 1233 66.4 376 4 US-09-180-100-22  
36 1233 66.4 387 1 US-08-470-299-4  
37 1233 66.4 389 3 US-09-131-247-14  
38 1233 66.4 396 2 US-08-784-512-3  
39 1233 66.4 396 3 US-09-176-228-3  
40 1233 66.4 424 4 US-09-333-593A-8  
41 1233 66.4 446 3 US-08-397-411-7  
42 1233 66.4 449 1 US-08-458-516-13  
43 1233 66.4 459 1 US-08-157-101A-7  
44 1233 66.4 475 4 US-09-740-002-27  
45 1233 66.4 476 2 US-08-378-939-10

#### ALIGNMENTS

RESULT 1  
US-07-940-861-43  
; Sequence 43, Application US/07940861  
; Patent No. 5547853  
; GENERAL INFORMATION:  
; APPLICANT: BIOGEN, INC.  
; APPLICANT: WALLNER, Barbara P.  
; APPLICANT: MILLER, Glenn T.  
; APPLICANT: ROSA, Margaret D.  
; TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE  
; TITLE OF INVENTION: FUNCTION ASSOCIATED ANTIGEN 3  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Neave  
; STREET: 875 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10022-6250  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/940,861  
; FILING DATE: 21-OCT-1992  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/02050  
; FILING DATE: 12-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/667,971  
; FILING DATE: 12-MAR-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,967  
; FILING DATE: 07-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HALEY, James F., Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: B151C1P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)715-0600  
; TELEFAX: (212)715-0673  
; TELEX: 14-8367  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 347 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

REFERENCE/DOCKET NUMBER: B151CIP2  
TELECOMMUNICATION INFORMATION:

APPLICATION NUMBER: US 07/667,971  
FILING DATE: 12-MAR-1991